



1/102

SEQUENCE LISTING

<110> VERTEX PHARMACEUTICALS INC.

<120> METHODS AND COMPOSITIONS FOR RAPID DEVELOPMENT OF  
SCREENING ASSAYS

<130> VPI/02-143W02

<140> PCT/US03/28078

<141> 2003-09-05

<150> 60/408,297

<151> 2002-09-05

<160> 86

<170> PatentIn version 3.2

<210> 1

<211> 795

<212> DNA

<213> Escherichia coli

<220>

<221> CDS

<222> (1)..(795)

<223> B-lactamase

<400> 1

atg agt cac cca gaa acg ctg gtg aaa gta aaa gat gct gaa gat cag	48
Met Ser His Pro Glu Thr Leu Val Lys Val Lys Asp Ala Glu Asp Gln	
1 5 10 15	
ttg ggt gca cga gtg ggt tac atc gaa ctg gat ctc aac agc ggt aag	96
Leu Gly Ala Arg Val Gly Tyr Ile Glu Leu Asp Leu Asn Ser Gly Lys	
20 25 30	
atc ctt gag agt ttt cgc ccc gaa gaa cgt ttt cca atg atg agc act	144
Ile Leu Glu Ser Phe Arg Pro Glu Glu Arg Phe Pro Met Met Ser Thr	
35 40 45	
ttt aaa gtt ctg cta tgt ggc gcg gta tta tcc cgt gtt gac gcc ggg	192
Phe Lys Val Leu Leu Cys Gly Ala Val Leu Ser Arg Val Asp Ala Gly	
50 55 60	
caa gag caa ctc ggt cgc cgc ata cac tat tct cag aat gac ttg gtt	240
Gln Glu Gln Leu Gly Arg Arg Ile His Tyr Ser Gln Asn Asp Leu Val	
65 70 75 80	
gag tac tca cca gtc aca gaa aag cat ctt acg gat ggc atg aca gta	288
Glu Tyr Ser Pro Val Thr Glu Lys His Leu Thr Asp Gly Met Thr Val	
85 90 95	
aga gaa tta tgc agt gct gcc ata acc atg agt gat aac act gcg gcc	336
Arg Glu Leu Cys Ser Ala Ala Ile Thr Met Ser Asp Asn Thr Ala Ala	
100 105 110	

aac tta ctt ctg aca acg atc gga gga ccg aag gag cta acc gct ttt	384
Asn Leu Leu Leu Thr Thr Ile Gly Gly Pro Lys Glu Leu Thr Ala Phe	
115 120 125	
ttg cac aac atg ggg gat cat gta act cgc ctt gat cgt tgg gaa ccg	432
Leu His Asn Met Gly Asp His Val Thr Arg Leu Asp Arg Trp Glu Pro	
130 135 140	
gag ctg aat gaa gcc ata cca aac gac gag cgt gac acc acg atg cct	480
Glu Leu Asn Glu Ala Ile Pro Asn Asp Glu Arg Asp Thr Thr Met Pro	
145 150 155 160	
gca gca atg gca aca acg ttg cgc aaa cta tta act ggc gaa cta ctt	528
Ala Ala Met Ala Thr Thr Leu Arg Lys Leu Leu Thr Gly Glu Leu Leu	
165 170 175	
act cta gct tcc cgg caa caa tta ata gac tgg atg gag gcg gat aaa	576
Thr Leu Ala Ser Arg Gln Gln Leu Ile Asp Trp Met Glu Ala Asp Lys	
180 185 190	
gtt gca gga cca ctt ctg cgc tcg gcc ctt ccg gct ggc tgg ttt att	624
Val Ala Gly Pro Leu Leu Arg Ser Ala Leu Pro Ala Gly Trp Phe Ile	
195 200 205	
gct gat aaa tct gga gcc ggt gag cgt ggg tct cgc ggt atc att gca	672
Ala Asp Lys Ser Gly Ala Gly Glu Arg Gly Ser Arg Gly Ile Ile Ala	
210 215 220	
gca ctg ggg cca gat ggt aag ccc tcc cgt atc gta gtt atc tac acg	720
Ala Leu Gly Pro Asp Gly Lys Pro Ser Arg Ile Val Val Ile Tyr Thr	
225 230 235 240	
acg ggg agt cag gca act atg gat gaa cga aat aga cag atc gct gag	768
Thr Gly Ser Gln Ala Thr Met Asp Glu Arg Asn Arg Gln Ile Ala Glu	
245 250 255	
ata ggt gcc tca ctg att aag cat tgg	795
Ile Gly Ala Ser Leu Ile Lys His Trp	
260 265	

&lt;210&gt; 2

&lt;211&gt; 265

&lt;212&gt; PRT

&lt;213&gt; Escherichia coli

&lt;400&gt; 2

Met Ser His Pro Glu Thr Leu Val Lys Val Lys Asp Ala Glu Asp Gln	
1 5 10 15	
Leu Gly Ala Arg Val Gly Tyr Ile Glu Leu Asp Leu Asn Ser Gly Lys	
20 25 30	
Ile Leu Glu Ser Phe Arg Pro Glu Glu Arg Phe Pro Met Met Ser Thr	
35 40 45	
Phe Lys Val Leu Leu Cys Gly Ala Val Leu Ser Arg Val Asp Ala Gly	
50 55 60	

Gln Glu Gln Leu Gly Arg Arg Ile His Tyr Ser Gln Asn Asp Leu Val  
 65 70 75 80  
 Glu Tyr Ser Pro Val Thr Glu Lys His Leu Thr Asp Gly Met Thr Val  
 85 90 95  
 Arg Glu Leu Cys Ser Ala Ala Ile Thr Met Ser Asp Asn Thr Ala Ala  
 100 105 110  
 Asn Leu Leu Leu Thr Thr Ile Gly Gly Pro Lys Glu Leu Thr Ala Phe  
 115 120 125  
 Leu His Asn Met Gly Asp His Val Thr Arg Leu Asp Arg Trp Glu Pro  
 130 135 140  
 Glu Leu Asn Glu Ala Ile Pro Asn Asp Glu Arg Asp Thr Thr Met Pro  
 145 150 155 160  
 Ala Ala Met Ala Thr Thr Leu Arg Lys Leu Leu Thr Gly Glu Leu Leu  
 165 170 175  
 Thr Leu Ala Ser Arg Gln Gln Leu Ile Asp Trp Met Glu Ala Asp Lys  
 180 185 190  
 Val Ala Gly Pro Leu Leu Arg Ser Ala Leu Pro Ala Gly Trp Phe Ile  
 195 200 205  
 Ala Asp Lys Ser Gly Ala Gly Glu Arg Gly Ser Arg Gly Ile Ile Ala  
 210 215 220  
 Ala Leu Gly Pro Asp Gly Lys Pro Ser Arg Ile Val Val Ile Tyr Thr  
 225 230 235 240  
 Thr Gly Ser Gln Ala Thr Met Asp Glu Arg Asn Arg Gln Ile Ala Glu  
 245 250 255  
 Ile Gly Ala Ser Leu Ile Lys His Trp  
 260 265

<210> 3  
 <211> 858  
 <212> DNA  
 <213> Escherichia coli

<220>  
 <221> CDS  
 <222> (1)..(858)  
 <223> B-lactamase

<400> 3  
 atg aga att caa cat ttc cgt gtc gcc ctt att ccc ttt ttt gcg gca  
 Met Arg Ile Gln His Phe Arg Val Ala Leu Ile Pro Phe Phe Ala Ala  
 1 5 10 15

ttt tgc ctt cct gtt ttt ggt cac cca gaa acg ctg gtg aaa gta aaa	96
Phe Cys Leu Pro Val Phe Gly His Pro Glu Thr Leu Val Lys Val Lys	
20 25 30	
gat gct gaa gat cag ttg ggt gca cga gtg ggt tac atc gaa ctg gat	144
Asp Ala Glu Asp Gln Leu Gly Ala Arg Val Gly Tyr Ile Glu Leu Asp	
35 40 45	
ctc aac agc ggt aag atc ctt gag agt ttt cgc ccc gaa gaa cgt ttt	192
Leu Asn Ser Gly Lys Ile Leu Glu Ser Phe Arg Pro Glu Glu Arg Phe	
50 55 60	
cca atg atg agc act ttt aaa gtt ctg cta tgt ggc gcg gta tta tcc	240
Pro Met Met Ser Thr Phe Lys Val Leu Leu Cys Gly Ala Val Leu Ser	
65 70 75 80	
cgt gtt gac gcc ggg caa gag caa ctc ggt cgc cgc ata cac tat tct	288
Arg Val Asp Ala Gly Gln Glu Gln Leu Gly Arg Arg Ile His Tyr Ser	
85 90 95	
cag aat gac ttg gtt gag tac tca cca gtc aca gaa aag cat ctt acg	336
Gln Asn Asp Leu Val Glu Tyr Ser Pro Val Thr Glu Lys His Leu Thr	
100 105 110	
gat ggc atg aca gta aga gaa tta tgc agt gct gcc ata acc atg agt	384
Asp Gly Met Thr Val Arg Glu Leu Cys Ser Ala Ala Ile Thr Met Ser	
115 120 125	
gat aac act gcg gcc aac tta ctt ctg aca acg atc gga gga ccg aag	432
Asp Asn Thr Ala Ala Asn Leu Leu Leu Thr Thr Ile Gly Gly Pro Lys	
130 135 140	
gag cta acc gct ttt ttg cac aac atg ggg gat cat gta act cgc ctt	480
Glu Leu Thr Ala Phe Leu His Asn Met Gly Asp His Val Thr Arg Leu	
145 150 155 160	
gat cgt tgg gaa ccg gag ctg aat gaa gcc ata cca aac gac gag cgt	528
Asp Arg Trp Glu Pro Glu Leu Asn Glu Ala Ile Pro Asn Asp Glu Arg	
165 170 175	
gac acc acg atg cct gca gca atg gca aca acg ttg cgc aaa cta tta	576
Asp Thr Thr Met Pro Ala Ala Met Ala Thr Thr Leu Arg Lys Leu Leu	
180 185 190	
act ggc gaa cta ctt act cta gct tcc cgg caa caa tta ata gac tgg	624
Thr Gly Glu Leu Leu Thr Leu Ala Ser Arg Gln Gln Leu Ile Asp Trp	
195 200 205	
atg gag gcg gat aaa gtt gca gga cca ctt ctg cgc tcg gcc ctt ccg	672
Met Glu Ala Asp Lys Val Ala Gly Pro Leu Leu Arg Ser Ala Leu Pro	
210 215 220	
gct ggc tgg ttt att gct gat aaa tct gga gcc ggt gag cgt ggg tct	720
Ala Gly Trp Phe Ile Ala Asp Lys Ser Gly Ala Gly Glu Arg Gly Ser	
225 230 235 240	

cgc ggt atc att gca gca ctg ggg cca gat ggt aag ccc tcc cgt atc 768  
 Arg Gly Ile Ile Ala Ala Leu Gly Pro Asp Gly Lys Pro Ser Arg Ile  
                   245                  250                  255

gta gtt atc tac acg acg ggg agt cag gca act atg gat gaa cga aat 816  
 Val Val Ile Tyr Thr Thr Gly Ser Gln Ala Thr Met Asp Glu Arg Asn  
                   260                  265                  270

aga cag atc gct gag ata ggt gcc tca ctg att aag cat tgg 858  
 Arg Gln Ile Ala Glu Ile Gly Ala Ser Leu Ile Lys His Trp  
                   275                  280                  285

<210> 4

<211> 286

<212> PRT

<213> Escherichia coli

<400> 4

Met Arg Ile Gln His Phe Arg Val Ala Leu Ile Pro Phe Phe Ala Ala  
 1                  5                  10                  15

Phe Cys Leu Pro Val Phe Gly His Pro Glu Thr Leu Val Lys Val Lys  
                   20                  25                  30

Asp Ala Glu Asp Gln Leu Gly Ala Arg Val Gly Tyr Ile Glu Leu Asp  
                   35                  40                  45

Leu Asn Ser Gly Lys Ile Leu Glu Ser Phe Arg Pro Glu Glu Arg Phe  
                   50                  55                  60

Pro Met Met Ser Thr Phe Lys Val Leu Leu Cys Gly Ala Val Leu Ser  
 65                  70                  75                  80

Arg Val Asp Ala Gly Gln Glu Gln Leu Gly Arg Arg Ile His Tyr Ser  
                   85                  90                  95

Gln Asn Asp Leu Val Glu Tyr Ser Pro Val Thr Glu Lys His Leu Thr  
                   100                  105                  110

Asp Gly Met Thr Val Arg Glu Leu Cys Ser Ala Ala Ile Thr Met Ser  
                   115                  120                  125

Asp Asn Thr Ala Ala Asn Leu Leu Leu Thr Thr Ile Gly Gly Pro Lys  
                   130                  135                  140

Glu Leu Thr Ala Phe Leu His Asn Met Gly Asp His Val Thr Arg Leu  
 145                  150                  155                  160

Asp Arg Trp Glu Pro Glu Leu Asn Glu Ala Ile Pro Asn Asp Glu Arg  
                   165                  170                  175

Asp Thr Thr Met Pro Ala Ala Met Ala Thr Thr Leu Arg Lys Leu Leu  
                   180                  185                  190

Thr Gly Glu Leu Leu Thr Leu Ala Ser Arg Gln Gln Leu Ile Asp Trp  
                   195                  200                  205

Met	Glu	Ala	Asp	Lys	Val	Ala	Gly	Pro	Leu	Leu	Arg	Ser	Ala	Leu	Pro
	210					215					220				
Ala	Gly	Trp	Phe	Ile	Ala	Asp	Lys	Ser	Gly	Ala	Gly	Glu	Arg	Gly	Ser
225					230					235					240
Arg	Gly	Ile	Ile	Ala	Ala	Leu	Gly	Pro	Asp	Gly	Lys	Pro	Ser	Arg	Ile
				245					250					255	
Val	Val	Ile	Tyr	Thr	Thr	Gly	Ser	Gln	Ala	Thr	Met	Asp	Glu	Arg	Asn
			260					265					270		
Arg	Gln	Ile	Ala	Glu	Ile	Gly	Ala	Ser	Leu	Ile	Lys	His	Trp		
	275						280					285			

&lt;210&gt; 5

&lt;211&gt; 795

&lt;212&gt; DNA

&lt;213&gt; Escherichia coli

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(795)

&lt;223&gt; B-lactamase

&lt;400&gt; 5

atg	ggg	cac	cca	gaa	acg	ctg	gtg	aaa	gta	aaa	gat	gct	gaa	gat	cag	48
Met	Gly	His	Pro	Glu	Thr	Leu	Val	Lys	Val	Lys	Asp	Ala	Glu	Asp	Gln	
1				5					10				15			
ttg	ggt	gca	cga	gtg	ggt	tac	atc	gaa	ctg	gat	ctc	aac	agc	ggt	aag	96
Leu	Gly	Ala	Arg	Val	Gly	Tyr	Ile	Glu	Leu	Asp	Leu	Asn	Ser	Gly	Lys	
			20					25					30			
atc	ctt	gag	agt	ttt	cgc	ccc	gaa	gaa	cgt	ttt	cca	atg	atg	agc	act	144
Ile	Leu	Glu	Ser	Phe	Arg	Pro	Glu	Glu	Arg	Phe	Pro	Met	Met	Ser	Thr	
			35				40					45				
ttt	aaa	gtt	ctg	cta	tgt	ggc	gcg	gta	tta	tcc	cgt	gat	gac	gcc	ggg	192
Phe	Lys	Val	Leu	Leu	Cys	Gly	Ala	Val	Leu	Ser	Arg	Asp	Asp	Ala	Gly	
	50					55					60					
caa	gag	caa	ctc	ggt	cgc	cgc	ata	cac	tat	tct	cag	aac	gac	ttg	gtt	240
Gln	Glu	Gln	Leu	Gly	Arg	Arg	Ile	His	Tyr	Ser	Gln	Asn	Asp	Leu	Val	
65					70				75					80		
gag	tac	tca	cca	gtc	aca	gaa	aag	cat	ctt	acg	gat	ggc	atg	aca	gta	288
Glu	Tyr	Ser	Pro	Val	Thr	Glu	Lys	His	Leu	Thr	Asp	Gly	Met	Thr	Val	
				85					90					95		
aga	gaa	tta	tgc	agt	gct	gcc	ata	acc	atg	agt	gat	aac	act	gcg	gcc	336
Arg	Glu	Leu	Cys	Ser	Ala	Ala	Ile	Thr	Met	Ser	Asp	Asn	Thr	Ala	Ala	
			100					105					110			
aac	tta	ctt	ctg	aca	acg	atc	gga	gga	ccg	aag	gag	cta	acc	gct	ttt	384
Asn	Leu	Leu	Leu	Thr	Thr	Ile	Gly	Gly	Pro	Lys	Glu	Leu	Thr	Ala	Phe	
		115					120					125				

ttg cac aac atg ggg gat cat gta act cgc ctt gat cat tgg gaa ccg	432
Leu His Asn Met Gly Asp His Val Thr Arg Leu Asp His Trp Glu Pro	
130 135 140	
gag ctg aat gaa gcc ata cca aac gac gag cgt gac acc acg atg cct	480
Glu Leu Asn Glu Ala Ile Pro Asn Asp Glu Arg Asp Thr Thr Met Pro	
145 150 155 160	
gta gca atg gca aca acg ttg cgc aaa cta tta act ggc gaa cta ctt	528
Val Ala Met Ala Thr Thr Leu Arg Lys Leu Leu Thr Gly Glu Leu Leu	
165 170 175	
act cta gct tcc cgg caa caa tta ata gac tgg atg gag gcg gat aaa	576
Thr Leu Ala Ser Arg Gln Gln Leu Ile Asp Trp Met Glu Ala Asp Lys	
180 185 190	
gtt gca gga cca ctt ctg cgc tcg gcc ctt ccg gct ggc tgg ttt att	624
Val Ala Gly Pro Leu Leu Arg Ser Ala Leu Pro Ala Gly Trp Phe Ile	
195 200 205	
gct gat aaa tct gga gcc ggt gag cgt ggg tct cgc ggt atc att gca	672
Ala Asp Lys Ser Gly Ala Gly Glu Arg Gly Ser Arg Gly Ile Ile Ala	
210 215 220	
gca ctg ggg cca gat ggt aag ccc tcc cgt atc gta gtt atc tac acg	720
Ala Leu Gly Pro Asp Gly Lys Pro Ser Arg Ile Val Val Ile Tyr Thr	
225 230 235 240	
acg ggg agt cag gca act atg gat gaa cga aat aga cag atc gct gag	768
Thr Gly Ser Gln Ala Thr Met Asp Glu Arg Asn Arg Gln Ile Ala Glu	
245 250 255	
ata ggt gcc tca ctg att aag cat tgg	795
Ile Gly Ala Ser Leu Ile Lys His Trp	
260 265	
<210> 6	
<211> 265	
<212> PRT	
<213> Escherichia coli	
<400> 6	
Met Gly His Pro Glu Thr Leu Val Lys Val Lys Asp Ala Glu Asp Gln	
1 5 10 15	
Leu Gly Ala Arg Val Gly Tyr Ile Glu Leu Asp Leu Asn Ser Gly Lys	
20 25 30	
Ile Leu Glu Ser Phe Arg Pro Glu Glu Arg Phe Pro Met Met Ser Thr	
35 40 45	
Phe Lys Val Leu Leu Cys Gly Ala Val Leu Ser Arg Asp Asp Ala Gly	
50 55 60	
Gln Glu Gln Leu Gly Arg Arg Ile His Tyr Ser Gln Asn Asp Leu Val	
65 70 75 80	

Glu	Tyr	Ser	Pro	Val	Thr	Glu	Lys	His	Leu	Thr	Asp	Gly	Met	Thr	Val	
				85					90					95		
Arg	Glu	Leu	Cys	Ser	Ala	Ala	Ile	Thr	Met	Ser	Asp	Asn	Thr	Ala	Ala	
				100					105					110		
Asn	Leu	Leu	Leu	Thr	Thr	Ile	Gly	Gly	Pro	Lys	Glu	Leu	Thr	Ala	Phe	
				115					120					125		
Leu	His	Asn	Met	Gly	Asp	His	Val	Thr	Arg	Leu	Asp	His	Trp	Glu	Pro	
				130					135					140		
Glu	Leu	Asn	Glu	Ala	Ile	Pro	Asn	Asp	Glu	Arg	Asp	Thr	Thr	Met	Pro	
				145					150					155		
Val	Ala	Met	Ala	Thr	Thr	Leu	Arg	Lys	Leu	Leu	Thr	Gly	Glu	Leu	Leu	
				165					170					175		
Thr	Leu	Ala	Ser	Arg	Gln	Gln	Leu	Ile	Asp	Trp	Met	Glu	Ala	Asp	Lys	
				180					185					190		
Val	Ala	Gly	Pro	Leu	Leu	Arg	Ser	Ala	Leu	Pro	Ala	Gly	Trp	Phe	Ile	
				195					200					205		
Ala	Asp	Lys	Ser	Gly	Ala	Gly	Glu	Arg	Gly	Ser	Arg	Gly	Ile	Ile	Ala	
				210					215					220		
Ala	Leu	Gly	Pro	Asp	Gly	Lys	Pro	Ser	Arg	Ile	Val	Val	Ile	Tyr	Thr	
				225					230					235		
Thr	Gly	Ser	Gln	Ala	Thr	Met	Asp	Glu	Arg	Asn	Arg	Gln	Ile	Ala	Glu	
				245					250					255		
Ile	Gly	Ala	Ser	Leu	Ile	Lys	His	Trp								
				260					265							

```
<210> 7
<211> 792
<212> DNA
<213> Escherichia coli
```

```
<220>
<221> CDS
<222> (1)..(792)
<223> B-lactamase
```

<400> 7																
atg	gac	cca	gaa	acg	ctg	gtg	aaa	gta	aaa	gat	gct	gaa	gat	cag	ttg	48
Met	Asp	Pro	Glu	Thr	Leu	Val	Lys	Val	Lys	Asp	Ala	Glu	Asp	Gln	Leu	
1				5					10					15		
ggt	gca	cga	gtg	ggc	tac	atc	gaa	ctg	gat	ctc	aac	agc	ggc	aag	atc	96
Gly	Ala	Arg	Val	Gly	Tyr	Ile	Glu	Leu	Asp	Leu	Asn	Ser	Gly	Lys	Ile	
			20					25					30			



ctt	gag	agt	ttt	cgc	ccc	gaa	gaa	cgt	ttt	cca	atg	atg	agc	act	ttt	144
Leu	Glu	Ser	Phe	Arg	Pro	Glu	Glu	Arg	Phe	Pro	Met	Met	Ser	Thr	Phe	
		35					40					45				
aaa	gtt	ctg	cta	tgt	ggc	gcg	gta	tta	tcc	cgt	att	gac	gcc	ggg	caa	192
Lys	Val	Leu	Leu	Cys	Gly	Ala	Val	Leu	Ser	Arg	Ile	Asp	Ala	Gly	Gln	
	50					55					60					
gag	caa	ctc	ggg	cgc	cgc	ata	cac	tat	tct	cag	aat	gac	ttg	gtt	gag	240
Glu	Gln	Leu	Gly	Arg	Arg	Ile	His	Tyr	Ser	Gln	Asn	Asp	Leu	Val	Glu	
65					70					75					80	
tac	tca	cca	gtc	aca	gaa	aag	cat	ctt	acg	gat	ggc	atg	aca	gta	aga	288
Tyr	Ser	Pro	Val	Thr	Glu	Lys	His	Leu	Thr	Asp	Gly	Met	Thr	Val	Arg	
				85					90					95		
gaa	tta	tgc	agt	gct	gcc	ata	acc	atg	agt	gat	aac	act	gcg	gcc	aac	336
Glu	Leu	Cys	Ser	Ala	Ala	Ile	Thr	Met	Ser	Asp	Asn	Thr	Ala	Ala	Asn	
			100					105					110			
tta	ctt	ctg	aca	acg	atc	gga	gga	ccg	aag	gag	cta	acc	gct	ttt	ttg	384
Leu	Leu	Leu	Thr	Thr	Ile	Gly	Gly	Pro	Lys	Glu	Leu	Thr	Ala	Phe	Leu	
		115				120						125				
cac	aac	atg	ggg	gat	cat	gta	act	cgc	ctt	gat	cat	tgg	gaa	ccg	gag	432
His	Asn	Met	Gly	Asp	His	Val	Thr	Arg	Leu	Asp	His	Trp	Glu	Pro	Glu	
	130					135					140					
ctg	aat	gaa	gcc	ata	cca	aac	gac	gag	cgt	gac	acc	acg	atg	cct	gta	480
Leu	Asn	Glu	Ala	Ile	Pro	Asn	Asp	Glu	Arg	Asp	Thr	Thr	Met	Pro	Val	
145					150					155					160	
gca	atg	gca	aca	acg	ttg	cgc	aaa	cta	tta	act	ggc	gaa	cta	ctt	act	528
Ala	Met	Ala	Thr		Leu	Arg	Lys	Leu	Leu	Thr	Gly	Glu	Leu	Leu	Thr	
				165					170					175		
cta	gct	tcc	cgg	caa	caa	tta	ata	gac	tgg	atg	gag	gcg	gat	aaa	gtt	576
Leu	Ala	Ser	Arg	Gln	Gln	Leu	Ile	Asp	Trp	Met	Glu	Ala	Asp	Lys	Val	
			180					185					190			
gca	gga	cca	ctt	ctg	cgc	tcg	gcc	ctt	ccg	gct	ggc	tgg	ttt	att	gct	624
Ala	Gly	Pro	Leu	Leu	Arg	Ser	Ala	Leu	Pro	Ala	Gly	Trp	Phe	Ile	Ala	
		195					200					205				
gat	aaa	tct	gga	gcc	ggg	gag	cgt	ggg	tct	cgc	ggg	atc	att	gca	gca	672
Asp	Lys	Ser	Gly	Ala	Gly	Glu	Arg	Gly	Ser	Arg	Gly	Ile	Ile	Ala	Ala	
	210					215					220					
ctg	ggg	cca	gat	ggg	aag	ccc	tcc	cgt	atc	gta	gtt	atc	tac	acg	acg	720
Leu	Gly	Pro	Asp	Gly	Lys	Pro	Ser	Arg	Ile	Val	Val	Ile	Tyr	Thr	Thr	
225					230					235					240	
ggg	agt	cag	gca	act	atg	gat	gaa	cga	aat	aga	cag	atc	gct	gag	ata	768
Gly	Ser	Gln	Ala	Thr	Met	Asp	Glu	Arg	Asn	Arg	Gln	Ile	Ala	Glu	Ile	
				245					250					255		

ggt gcc tca ctg att aag cat tgg  
 Gly Ala Ser Leu Ile Lys His Trp  
 260

<210> 8  
 <211> 264  
 <212> PRT  
 <213> Escherichia coli

<400> 8  
 Met Asp Pro Glu Thr Leu Val Lys Val Lys Asp Ala Glu Asp Gln Leu  
 1 5 10 15  
 Gly Ala Arg Val Gly Tyr Ile Glu Leu Asp Leu Asn Ser Gly Lys Ile  
 20 25 30  
 Leu Glu Ser Phe Arg Pro Glu Glu Arg Phe Pro Met Met Ser Thr Phe  
 35 40 45  
 Lys Val Leu Leu Cys Gly Ala Val Leu Ser Arg Ile Asp Ala Gly Gln  
 50 55 60  
 Glu Gln Leu Gly Arg Arg Ile His Tyr Ser Gln Asn Asp Leu Val Glu  
 65 70 75 80  
 Tyr Ser Pro Val Thr Glu Lys His Leu Thr Asp Gly Met Thr Val Arg  
 85 90 95  
 Glu Leu Cys Ser Ala Ala Ile Thr Met Ser Asp Asn Thr Ala Ala Asn  
 100 105 110  
 Leu Leu Leu Thr Thr Ile Gly Gly Pro Lys Glu Leu Thr Ala Phe Leu  
 115 120 125  
 His Asn Met Gly Asp His Val Thr Arg Leu Asp His Trp Glu Pro Glu  
 130 135 140  
 Leu Asn Glu Ala Ile Pro Asn Asp Glu Arg Asp Thr Thr Met Pro Val  
 145 150 155 160  
 Ala Met Ala Thr Thr Leu Arg Lys Leu Leu Thr Gly Glu Leu Leu Thr  
 165 170 175  
 Leu Ala Ser Arg Gln Gln Leu Ile Asp Trp Met Glu Ala Asp Lys Val  
 180 185 190  
 Ala Gly Pro Leu Leu Arg Ser Ala Leu Pro Ala Gly Trp Phe Ile Ala  
 195 200 205  
 Asp Lys Ser Gly Ala Gly Glu Arg Gly Ser Arg Gly Ile Ile Ala Ala  
 210 215 220  
 Leu Gly Pro Asp Gly Lys Pro Ser Arg Ile Val Val Ile Tyr Thr Thr  
 225 230 235 240  
 Gly Ser Gln Ala Thr Met Asp Glu Arg Asn Arg Gln Ile Ala Glu Ile  
 245 250 255

Gly Ala Ser Leu Ile Lys His Trp  
260

<210> 9  
<211> 786  
<212> DNA  
<213> Escherichia coli

<220>  
<221> CDS  
<222> (1)..(786)  
<223> B-lactamase

<400> 9  
atg aaa gat gat ttt gca aaa ctt gag gaa caa ttt gat gca aaa ctc 48  
Met Lys Asp Asp Phe Ala Lys Leu Glu Glu Gln Phe Asp Ala Lys Leu  
1 5 10 15  
ggg atc ttt gca ttg gat aca ggt aca aac cgg acg gta gcg tat cgg 96  
Gly Ile Phe Ala Leu Asp Thr Gly Thr Asn Arg Thr Val Ala Tyr Arg  
20 25 30  
ccg gat gag cgt ttt gct ttt gct tcg acg att aag gct tta act gta 144  
Pro Asp Glu Arg Phe Ala Phe Ala Ser Thr Ile Lys Ala Leu Thr Val  
35 40 45  
ggc gtg ctt ttg caa cag aaa tca ata gaa gat ctg aac cag aga ata 192  
Gly Val Leu Leu Gln Gln Lys Ser Ile Glu Asp Leu Asn Gln Arg Ile  
50 55 60  
aca tat aca cgt gat gat ctt gta aac tac aac ccg att acg gaa aag 240  
Thr Tyr Thr Arg Asp Asp Leu Val Asn Tyr Asn Pro Ile Thr Glu Lys  
65 70 75 80  
cac gtt gat acg gga atg acg ctc aaa gag ctt gcg gat gct tcg ctt 288  
His Val Asp Thr Gly Met Thr Leu Lys Glu Leu Ala Asp Ala Ser Leu  
85 90 95  
cga tat agt gac aat gcg gca cag aat ctc att ctt aaa caa att ggc 336  
Arg Tyr Ser Asp Asn Ala Ala Gln Asn Leu Ile Leu Lys Gln Ile Gly  
100 105 110  
gga cct gaa agt ttg aaa aag gaa ctg agg aag att ggt gat gag gtt 384  
Gly Pro Glu Ser Leu Lys Lys Glu Leu Arg Lys Ile Gly Asp Glu Val  
115 120 125  
aca aat ccc gaa cga ttc gaa cca gag tta aat gaa gtg aat ccg ggt 432  
Thr Asn Pro Glu Arg Phe Glu Pro Glu Leu Asn Glu Val Asn Pro Gly  
130 135 140  
gaa act cag gat acc agt aca gca aga gca ctt gtc aca agc ctt cga 480  
Glu Thr Gln Asp Thr Ser Thr Ala Arg Ala Leu Val Thr Ser Leu Arg  
145 150 155 160

12/102

gcc ttt gct ctt gaa gat aaa ctt cca agt gaa aaa cgc gag ctt tta	528
Ala Phe Ala Leu Glu Asp Lys Leu Pro Ser Glu Lys Arg Glu Leu Leu	
165 170 175	
atc gat tgg atg aaa cga aat acc act gga gac gcc tta atc cgt gcc	576
Ile Asp Trp Met Lys Arg Asn Thr Thr Gly Asp Ala Leu Ile Arg Ala	
180 185 190	
gga gcg gca tca tat gga acc cgg aat gac att gcc atc att tgg ccg	624
Gly Ala Ala Ser Tyr Gly Thr Arg Asn Asp Ile Ala Ile Ile Trp Pro	
195 200 205	
cca aaa gga gat cct gtc ggt gtg ccg gac ggt tgg gaa gtg gct gat	672
Pro Lys Gly Asp Pro Val Gly Val Pro Asp Gly Trp Glu Val Ala Asp	
210 215 220	
aaa act gtt ctt gca gta tta tcc agc agg gat aaa aag gac gcc aag	720
Lys Thr Val Leu Ala Val Leu Ser Ser Arg Asp Lys Lys Asp Ala Lys	
225 230 235 240	
tat gat gat aaa ctt att gca gag gca aca aag gtg gta atg aaa gcc	768
Tyr Asp Asp Lys Leu Ile Ala Glu Ala Thr Lys Val Val Met Lys Ala	
245 250 255	
tta aac atg aac ggc aaa	786
Leu Asn Met Asn Gly Lys	
260	

<210> 10

<211> 262

<212> PRT

<213> Escherichia coli )

<400> 10

Met Lys Asp Asp Phe Ala Lys Leu Glu Glu Gln Phe Asp Ala Lys Leu	
1 5 10 15	
Gly Ile Phe Ala Leu Asp Thr Gly Thr Asn Arg Thr Val Ala Tyr Arg	
20 25 30	
Pro Asp Glu Arg Phe Ala Phe Ala Ser Thr Ile Lys Ala Leu Thr Val	
35 40 45	
Gly Val Leu Leu Gln Gln Lys Ser Ile Glu Asp Leu Asn Gln Arg Ile	
50 55 60	
Thr Tyr Thr Arg Asp Asp Leu Val Asn Tyr Asn Pro Ile Thr Glu Lys	
65 70 75 80	
His Val Asp Thr Gly Met Thr Leu Lys Glu Leu Ala Asp Ala Ser Leu	
85 90 95	
Arg Tyr Ser Asp Asn Ala Ala Gln Asn Leu Ile Leu Lys Gln Ile Gly	
100 105 110	
Gly Pro Glu Ser Leu Lys Lys Glu Leu Arg Lys Ile Gly Asp Glu Val	
115 120 125	

Thr Asn Pro Glu Arg Phe Glu Pro Glu Leu Asn Glu Val Asn Pro Gly  
 130 135 140  
 Glu Thr Gln Asp Thr Ser Thr Ala Arg Ala Leu Val Thr Ser Leu Arg  
 145 150 155 160  
 Ala Phe Ala Leu Glu Asp Lys Leu Pro Ser Glu Lys Arg Glu Leu Leu  
 165 170 175  
 Ile Asp Trp Met Lys Arg Asn Thr Thr Gly Asp Ala Leu Ile Arg Ala  
 180 185 190  
 Gly Ala Ala Ser Tyr Gly Thr Arg Asn Asp Ile Ala Ile Ile Trp Pro  
 195 200 205  
 Pro Lys Gly Asp Pro Val Gly Val Pro Asp Gly Trp Glu Val Ala Asp  
 210 215 220  
 Lys Thr Val Leu Ala Val Leu Ser Ser Arg Asp Lys Lys Asp Ala Lys  
 225 230 235 240  
 Tyr Asp Asp Lys Leu Ile Ala Glu Ala Thr Lys Val Val Met Lys Ala  
 245 250 255  
 Leu Asn Met Asn Gly Lys  
 260

<210> 11  
 <211> 720  
 <212> DNA  
 <213> *Aequorea victoria*

<220>  
 <221> CDS  
 <222> (1)..(720)  
 <223> mutant green fluorescent protein

<400> 11  
 atg gtg agc aag ggc gag gag ctg ttc acc ggg gtg gtg ccc atc ctg 48  
 Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu  
 1 5 10 15  
 gtc gag ctg gac ggc gac gta aac ggc cac aag ttc agc gtg tcc ggc 96  
 Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly  
 20 25 30  
 gag ggc gag ggc gat gcc acc tac ggc aag ctg acc ctg aag ttc atc 144  
 Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile  
 35 40 45  
 tgc acc acc ggc aag ctg ccc gtg ccc tgg ccc acc ctc gtg acc acc 192  
 Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr  
 50 55 60

14/102

ttc tcc tac ggc gtg cag tgc ttc agc cgc tac ccc gac cac atg aag 240  
Phe Ser Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys  
65 70 75 80

cag cac gac ttc ttc aag tcc gcc atg ccc gaa ggc tac gtc cag gag 288  
Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu  
85 90 95

cgc acc atc ttc ttc aag gac gac ggc aac tac aag acc cgc gcc gag 336  
Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu  
100 105 110

gtg aag ttc gag ggc gac acc ctg gtg aac cgc atc gag ctg aag ggc 384  
Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly  
115 120 125

atc gac ttc aag gag gac ggc aac atc ctg ggg cac aac ctg gag tac 432  
Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Asn Leu Glu Tyr  
130 135 140

aac tac aac agc cac aac gtc tat atc atg gcc gac aag cag aag aac 480  
Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn  
145 150 155 160

ggc atc aag gtg aac ttc aag atc cgc cac aac atc gag gac ggc agc 528  
Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser  
165 170 175

gtg cag ctc gcc gac cac tac cag cag aac acc ccc atc ggc gac ggc 576  
Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly  
180 185 190

ccc gtg ctg ctg ccc gac aac cac tac ctg agc acc cag tcc gcc ctg 624  
Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu  
195 200 205

agc aaa gac ccc aac gag aag cgc gat cac atg gtc ctg ctg gag ttc 672  
Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe  
210 215 220

gtg acc gcc gcc ggg atc act ctc ggc atg gac gag ctg tac aag taa 720  
Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys  
225 230 235

<210> 12

<211> 239

<212> PRT

<213> Aequorea victoria

<400> 12

Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu  
1 5 10 15

Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly  
20 25 30

15/102

Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile  
35 40 45  
Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr  
50 55 60  
Phe Ser Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys  
65 70 75 80  
Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu  
85 90 95  
Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu  
100 105 110  
Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly  
115 120 125  
Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Asn Leu Glu Tyr  
130 135 140  
Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn  
145 150 155 160  
Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser  
165 170 175  
Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly  
180 185 190  
Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu  
195 200 205  
Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe  
210 215 220  
Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys  
225 230 235

<210> 13  
<211> 690  
<212> DNA  
<213> Anemonia majano

<220>  
<221> CDS  
<222> (1)..(690)  
<223> fluorescent protein

<400> 13  
atg gct ctt tca aac aag ttt atc gga gat gac atg aaa atg acc tac 48  
Met Ala Leu Ser Asn Lys Phe Ile Gly Asp Asp Met Lys Met Thr Tyr  
1 5 10 15  
cat atg gat ggc tgt gtc aat ggg cat tac ttt acc gtc aaa ggt gaa 96  
His Met Asp Gly Cys Val Asn Gly His Tyr Phe Thr Val Lys Gly Glu  
20 25 30

ggc aac ggg aag cca tac gaa ggg acg cag act tcg act ttt aaa gtc Gly Asn Gly Lys Pro Tyr Glu Gly Thr Gln Thr Ser Thr Phe Lys Val 35 40 45	144
acc atg gcc aac ggt ggg ccc ctt gca ttc tcc ttt gac ata cta tct Thr Met Ala Asn Gly Gly Pro Leu Ala Phe Ser Phe Asp Ile Leu Ser 50 55 60	192
aca gtg ttc aaa tat gga aat cga tgc ttt act gcg tat cct acc agt Thr Val Phe Lys Tyr Gly Asn Arg Cys Phe Thr Ala Tyr Pro Thr Ser 65 70 75 80	240
atg ccc gac tat ttc aaa caa gca ttt cct gac gga atg tca tat gaa Met Pro Asp Tyr Phe Lys Gln Ala Phe Pro Asp Gly Met Ser Tyr Glu 85 90 95	288
agg act ttt acc tat gaa gat gga gga gtt gct aca gcc agt tgg gaa Arg Thr Phe Thr Tyr Glu Asp Gly Gly Val Ala Thr Ala Ser Trp Glu 100 105 110	336
ata agc ctt aaa ggc aac tgc ttt gag cac aaa tcc acg ttt cat gga Ile Ser Leu Lys Gly Asn Cys Phe Glu His Lys Ser Thr Phe His Gly 115 120 125	384
gtg aac ttt cct gct gat gga cct gtg atg gcg aag aag aca act ggt Val Asn Phe Pro Ala Asp Gly Pro Val Met Ala Lys Lys Thr Thr Gly 130 135 140	432
tgg gac cca tct ttt gag aaa atg act gtc tgc gat gga ata ttg aag Trp Asp Pro Ser Phe Glu Lys Met Thr Val Cys Asp Gly Ile Leu Lys 145 150 155 160	480
ggg gat gtc acc gcg ttc ctc atg ctg caa gga ggt ggc aat tac aga Gly Asp Val Thr Ala Phe Leu Met Leu Gln Gly Gly Gly Asn Tyr Arg 165 170 175	528
tgc caa ttc cac act tct tac aag aca aaa aaa ccg gtg acg atg cca Cys Gln Phe His Thr Ser Tyr Lys Thr Lys Lys Pro Val Thr Met Pro 180 185 190	576
cca aac cat gtg gtg gaa cat cgc att gcg agg acc gac ctt gac aaa Pro Asn His Val Val Glu His Arg Ile Ala Arg Thr Asp Leu Asp Lys 195 200 205	624
ggg ggc aac agt gtt cag ctg acg gag cac gct gtt gca cat ata acc Gly Gly Asn Ser Val Gln Leu Thr Glu His Ala Val Ala His Ile Thr 210 215 220	672
tct gtt gtc cct ttc tga Ser Val Val Pro Phe 225	690

&lt;210&gt; 14

&lt;211&gt; 229

&lt;212&gt; PRT

&lt;213&gt; Anemonia majano



&lt;400&gt; 14

Met Ala Leu Ser Asn Lys Phe Ile Gly Asp Asp Met Lys Met Thr Tyr  
 1 5 10 15

His Met Asp Gly Cys Val Asn Gly His Tyr Phe Thr Val Lys Gly Glu  
 20 25 30

Gly Asn Gly Lys Pro Tyr Glu Gly Thr Gln Thr Ser Thr Phe Lys Val  
 35 40 45

Thr Met Ala Asn Gly Gly Pro Leu Ala Phe Ser Phe Asp Ile Leu Ser  
 50 55 60

Thr Val Phe Lys Tyr Gly Asn Arg Cys Phe Thr Ala Tyr Pro Thr Ser  
 65 70 75 80

Met Pro Asp Tyr Phe Lys Gln Ala Phe Pro Asp Gly Met Ser Tyr Glu  
 85 90 95

Arg Thr Phe Thr Tyr Glu Asp Gly Gly Val Ala Thr Ala Ser Trp Glu  
 100 105 110

Ile Ser Leu Lys Gly Asn Cys Phe Glu His Lys Ser Thr Phe His Gly  
 115 120 125

Val Asn Phe Pro Ala Asp Gly Pro Val Met Ala Lys Lys Thr Thr Gly  
 130 135 140

Trp Asp Pro Ser Phe Glu Lys Met Thr Val Cys Asp Gly Ile Leu Lys  
 145 150 155 160

Gly Asp Val Thr Ala Phe Leu Met Leu Gln Gly Gly Gly Asn Tyr Arg  
 165 170 175

Cys Gln Phe His Thr Ser Tyr Lys Thr Lys Lys Pro Val Thr Met Pro  
 180 185 190

Pro Asn His Val Val Glu His Arg Ile Ala Arg Thr Asp Leu Asp Lys  
 195 200 205

Gly Gly Asn Ser Val Gln Leu Thr Glu His Ala Val Ala His Ile Thr  
 210 215 220

Ser Val Val Pro Phe  
 225

&lt;210&gt; 15

&lt;211&gt; 696

&lt;212&gt; DNA

&lt;213&gt; Zoanthus sp.

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(696)

&lt;223&gt; fluorescent protein

&lt;400&gt; 15

atg gct cag tca aag cac ggt cta aca aaa gaa atg aca atg aaa tac	48
Met Ala Gln Ser Lys His Gly Leu Thr Lys Glu Met Thr Met Lys Tyr	
1 5 10 15	
cgt atg gaa ggg tgc gtc gat gga cat aaa ttt gtg atc acg gga gag	96
Arg Met Glu Gly Cys Val Asp Gly His Lys Phe Val Ile Thr Gly Glu	
20 25 30	
ggc att gga tat ccg ttc aaa ggg aaa cag gct att aat ctg tgt gtg	144
Gly Ile Gly Tyr Pro Phe Lys Gly Lys Gln Ala Ile Asn Leu Cys Val	
35 40 45	
gtc gaa ggt gga cca ttg cca ttt gcc gaa gac ata ttg tca gct gcc	192
Val Glu Gly Gly Pro Leu Pro Phe Ala Glu Asp Ile Leu Ser Ala Ala	
50 55 60	
ttt aac tac gga aac agg gtt ttc act gaa tat cct caa gac ata gtt	240
Phe Asn Tyr Gly Asn Arg Val Phe Thr Glu Tyr Pro Gln Asp Ile Val	
65 70 75 80	
gac tat ttc aag aac tcg tgt cct gct gga tat aca tgg gac agg tct	288
Asp Tyr Phe Lys Asn Ser Cys Pro Ala Gly Tyr Thr Trp Asp Arg Ser	
85 90 95	
ttt ctc ttt gag gat gga gca gtt tgc ata tgt aat gca gat ata aca	336
Phe Leu Phe Glu Asp Gly Ala Val Cys Ile Cys Asn Ala Asp Ile Thr	
100 105 110	
gtg agt gtt gaa gaa aac tgc atg tat cat gag tcc aaa ttt tat gga	384
Val Ser Val Glu Glu Asn Cys Met Tyr His Glu Ser Lys Phe Tyr Gly	
115 120 125	
gtg aat ttt cct gct gat gga cct gtg atg aaa aag atg aca gat aac	432
Val Asn Phe Pro Ala Asp Gly Pro Val Met Lys Lys Met Thr Asp Asn	
130 135 140	
tgg gag cca tcc tgc gag aag atc ata cca gta cct aag cag ggg ata	480
Trp Glu Pro Ser Cys Glu Lys Ile Ile Pro Val Pro Lys Gln Gly Ile	
145 150 155 160	
ttg aaa ggg gat gtc tcc atg tac ctc ctt ctg aag gat ggt ggg cgt	528
Leu Lys Gly Asp Val Ser Met Tyr Leu Leu Leu Lys Asp Gly Gly Arg	
165 170 175	
tta cgg tgc caa ttc gac aca gtt tac aaa gca aag tct gtg cca aga	576
Leu Arg Cys Gln Phe Asp Thr Val Tyr Lys Ala Lys Ser Val Pro Arg	
180 185 190	
aag atg ccg gac tgg cac ttc atc cag cat aag ctc acc cgt gaa gac	624
Lys Met Pro Asp Trp His Phe Ile Gln His Lys Leu Thr Arg Glu Asp	
195 200 205	
cgc agc gat gct aag aat cag aaa tgg cat ctg aca gaa cat gct att	672
Arg Ser Asp Ala Lys Asn Gln Lys Trp His Leu Thr Glu His Ala Ile	
210 215 220	

gca tcc gga tct gca ttg ccc tga  
 Ala Ser Gly Ser Ala Leu Pro  
 225 230

<210> 16  
 <211> 231  
 <212> PRT  
 <213> Zoanthus sp.

<400> 16  
 Met Ala Gln Ser Lys His Gly Leu Thr Lys Glu Met Thr Met Lys Tyr  
 1 5 10 15  
 Arg Met Glu Gly Cys Val Asp Gly His Lys Phe Val Ile Thr Gly Glu  
 20 25 30  
 Gly Ile Gly Tyr Pro Phe Lys Gly Lys Gln Ala Ile Asn Leu Cys Val  
 35 40 45  
 Val Glu Gly Gly Pro Leu Pro Phe Ala Glu Asp Ile Leu Ser Ala Ala  
 50 55 60  
 Phe Asn Tyr Gly Asn Arg Val Phe Thr Glu Tyr Pro Gln Asp Ile Val  
 65 70 75 80  
 Asp Tyr Phe Lys Asn Ser Cys Pro Ala Gly Tyr Thr Trp Asp Arg Ser  
 85 90 95  
 Phe Leu Phe Glu Asp Gly Ala Val Cys Ile Cys Asn Ala Asp Ile Thr  
 100 105 110  
 Val Ser Val Glu Glu Asn Cys Met Tyr His Glu Ser Lys Phe Tyr Gly  
 115 120 125  
 Val Asn Phe Pro Ala Asp Gly Pro Val Met Lys Lys Met Thr Asp Asn  
 130 135 140  
 Trp Glu Pro Ser Cys Glu Lys Ile Ile Pro Val Pro Lys Gln Gly Ile  
 145 150 155 160  
 Leu Lys Gly Asp Val Ser Met Tyr Leu Leu Leu Lys Asp Gly Gly Arg  
 165 170 175  
 Leu Arg Cys Gln Phe Asp Thr Val Tyr Lys Ala Lys Ser Val Pro Arg  
 180 185 190  
 Lys Met Pro Asp Trp His Phe Ile Gln His Lys Leu Thr Arg Glu Asp  
 195 200 205  
 Arg Ser Asp Ala Lys Asn Gln Lys Trp His Leu Thr Glu His Ala Ile  
 210 215 220  
 Ala Ser Gly Ser Ala Leu Pro  
 225 230

```
<220>
<221> CDS
<222> (1)..(696)
<223> fluorescent protein
```

<400>	17																
atg	gct	cat	tca	aag	cac	ggt	cta	aaa	gaa	gaa	atg	aca	atg	aaa	tac	48	
Met	Ala	His	Ser	Lys	His	Gly	Leu	Lys	Glu	Glu	Met	Thr	Met	Lys	Tyr		
1			5					10						15			
cac	atg	gaa	ggg	tgc	gtc	aac	gga	cat	aaa	ttt	gtg	atc	acg	ggc	gaa	96	
His	Met	Glu	Gly	Cys	Val	Asn	Gly	His	Lys	Phe	Val	Ile	Thr	Gly	Glu		
			20					25					30				
ggc	att	gga	tat	ccg	ttc	aaa	ggg	aaa	cag	act	att	aat	ctg	tgt	gtg	144	
Gly	Ile	Gly	Tyr	Pro	Phe	Lys	Gly	Lys	Gln	Thr	Ile	Asn	Leu	Cys	Val		
		35					40					45					
atc	gaa	ggg	gga	cca	ttg	cca	ttt	tcc	gaa	gac	ata	ttg	tca	gct	ggc	192	
Ile	Glu	Gly	Gly	Pro	Leu	Pro	Phe	Ser	Glu	Asp	Ile	Leu	Ser	Ala	Gly		
	50					55					60						
ttt	aag	tac	gga	gac	agg	att	ttc	act	gaa	tat	cct	caa	gac	ata	gta	240	
Phe	Lys	Tyr	Gly	Asp	Arg	Ile	Phe	Thr	Glu	Tyr	Pro	Gln	Asp	Ile	Val		
65					70				75						80		
gac	tat	ttc	aag	aac	tcg	tgt	cct	gct	gga	tat	aca	tgg	ggc	agg	tct	288	
Asp	Tyr	Phe	Lys	Asn	Ser	Cys	Pro	Ala	Gly	Tyr	Thr	Trp	Gly	Arg	Ser		
			85						90					95			
ttt	ctc	ttt	gag	gat	gga	gca	gtc	tgc	ata	tgc	aat	gta	gat	ata	aca	336	
Phe	Leu	Phe	Glu	Asp	Gly	Ala	Val	Cys	Ile	Cys	Asn	Val	Asp	Ile	Thr		
			100					105					110				
gtg	agt	gtc	aaa	gaa	aac	tgc	att	tat	cat	aag	agc	ata	ttt	aat	gga	384	
Val	Ser	Val	Lys	Glu	Asn	Cys	Ile	Tyr	His	Lys	Ser	Ile	Phe	Asn	Gly		
		115					120					125					
atg	aat	ttt	cct	gct	gat	gga	cct	gtg	atg	aaa	aag	atg	aca	act	aac	432	
Met	Asn	Phe	Pro	Ala	Asp	Gly	Pro	Val	Met	Lys	Lys	Met	Thr	Thr	Asn		
	130					135					140						
tgg	gaa	gca	tcc	tgc	gag	aag	atc	atg	cca	gta	cct	aag	cag	ggg	ata	480	
Trp	Glu	Ala	Ser	Cys	Glu	Lys	Ile	Met	Pro	Val	Pro	Lys	Gln	Gly	Ile		
145					150				155						160		
ctg	aaa	ggg	gat	gtc	tcc	atg	tac	ctc	ctt	ctg	aag	gat	ggg	ggg	cgt	528	
Leu	Lys	Gly	Asp	Val	Ser	Met	Tyr	Leu	Leu	Leu	Lys	Asp	Gly	Gly	Arg		
				165					170					175			
tac	cgg	tgc	cag	ttc	gac	aca	gtt	tac	aaa	gca	aag	tct	gtg	cca	agt	576	
Tyr	Arg	Cys	Gln	Phe	Asp	Thr	Val	Tyr	Lys	Ala	Lys	Ser	Val	Pro	Ser		
			180					185					190				

aag atg ccg gag tgg cac ttc atc cag cat aag ctc ctc cgt gaa gac 624  
 Lys Met Pro Glu Trp His Phe Ile Gln His Lys Leu Arg Glu Asp  
 195 200 205

cgc agc gat gct aag aat cag aag tgg cag ctg aca gag cat gct att 672  
 Arg Ser Asp Ala Lys Asn Gln Lys Trp Gln Leu Thr Glu His Ala Ile  
 210 215 220

gca ttc cct tct gcc ttg gcc tga 696  
 Ala Phe Pro Ser Ala Leu Ala  
 225 230

<210> 18  
 <211> 231  
 <212> PRT  
 <213> Zoanthus sp.

<400> 18  
 Met Ala His Ser Lys His Gly Leu Lys Glu Glu Met Thr Met Lys Tyr  
 1 5 10 15

His Met Glu Gly Cys Val Asn Gly His Lys Phe Val Ile Thr Gly Glu  
 20 25 30

Gly Ile Gly Tyr Pro Phe Lys Gly Lys Gln Thr Ile Asn Leu Cys Val  
 35 40 45

Ile Glu Gly Gly Pro Leu Pro Phe Ser Glu Asp Ile Leu Ser Ala Gly  
 50 55 60

Phe Lys Tyr Gly Asp Arg Ile Phe Thr Glu Tyr Pro Gln Asp Ile Val  
 65 70 75 80

Asp Tyr Phe Lys Asn Ser Cys Pro Ala Gly Tyr Thr Trp Gly Arg Ser  
 85 90 95

Phe Leu Phe Glu Asp Gly Ala Val Cys Ile Cys Asn Val Asp Ile Thr  
 100 105 110

Val Ser Val Lys Glu Asn Cys Ile Tyr His Lys Ser Ile Phe Asn Gly  
 115 120 125

Met Asn Phe Pro Ala Asp Gly Pro Val Met Lys Lys Met Thr Thr Asn  
 130 135 140

Trp Glu Ala Ser Cys Glu Lys Ile Met Pro Val Pro Lys Gln Gly Ile  
 145 150 155 160

Leu Lys Gly Asp Val Ser Met Tyr Leu Leu Leu Lys Asp Gly Gly Arg  
 165 170 175

Tyr Arg Cys Gln Phe Asp Thr Val Tyr Lys Ala Lys Ser Val Pro Ser  
 180 185 190

Lys Met Pro Glu Trp His Phe Ile Gln His Lys Leu Leu Arg Glu Asp  
 195 200 205

Arg Ser Asp Ala Lys Asn Gln Lys Trp Gln Leu Thr Glu His Ala Ile  
 210 215 220

Ala Phe Pro Ser Ala Leu Ala  
 225 230

<210> 19

<211> 699

<212> DNA

<213> *Discosoma striata*

<220>

<221> CDS

<222> (1)..(699)

<223> fluorescent protein

<400> 19

atg agt tgt tcc aag agt gtg atc aag gaa gaa atg ttg atc gat ctt 48  
 Met Ser Cys Ser Lys Ser Val Ile Lys Glu Met Leu Ile Asp Leu  
 1 5 10 15

cat ctg gaa gga acg ttc aat ggg cac tac ttt gaa ata aaa ggc aaa 96  
 His Leu Glu Gly Thr Phe Asn Gly His Tyr Phe Glu Ile Lys Gly Lys  
 20 25 30

gga aaa gga cag cct aat gaa ggc acc aat acc gtc acg ctc gag gtt 144  
 Gly Lys Gly Gln Pro Asn Glu Gly Thr Asn Thr Val Thr Leu Glu Val  
 35 40 45

acc aag ggt gga cct ctg cca ttt ggt tgg cat att ttg tgc cca caa 192  
 Thr Lys Gly Gly Pro Leu Phe Gly Trp His Ile Leu Cys Pro Gln  
 50 55 60

ttt cag tat gga aac aag gca ttt gtc cac cac cct gac aac ata cat 240  
 Phe Gln Tyr Gly Asn Lys Ala Phe Val His His Pro Asp Asn Ile His  
 65 70 75 80

gat tat cta aag ctg tca ttt ccg gag gga tat aca tgg gaa cgg tcc 288  
 Asp Tyr Leu Lys Leu Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ser  
 85 90 95

atg cac ttt gaa gac ggt ggc ttg tgt tgt atc acc aat gat atc agt 336  
 Met His Phe Glu Asp Gly Gly Leu Cys Cys Ile Thr Asn Asp Ile Ser  
 100 105 110

ttg aca ggc aac tgt ttc tac tac gac atc aag ttc act ggc ttg aac 384  
 Leu Thr Gly Asn Cys Phe Tyr Tyr Asp Ile Lys Phe Thr Gly Leu Asn  
 115 120 125

ttt cct cca aat gga ccc gtt gtg cag aag aag aca act ggc tgg gaa 432  
 Phe Pro Pro Asn Gly Pro Val Val Gln Lys Lys Thr Thr Gly Trp Glu  
 130 135 140

ccg agc act gag cgt ttg tat cct cgt gat ggt gtg ttg ata gga gac 480  
 Pro Ser Thr Glu Arg Leu Tyr Pro Arg Asp Gly Val Leu Ile Gly Asp  
 145 150 155 160

atc cat cat gct ctg aca gtt gaa gga ggt ggt cat tac gca tgt gac 528  
 Ile His His Ala Leu Thr Val Glu Gly Gly Gly His Tyr Ala Cys Asp  
                   165                  170                  175

att aaa act gtt tac agg gcc aag aag gcc gcc ttg aag atg cca ggg 576  
 Ile Lys Thr Val Tyr Arg Ala Lys Lys Ala Ala Leu Lys Met Pro Gly  
                   180                  185                  190

tat cac tat gtt gac acc aaa ctg gtt ata tgg aac aac gac aaa gaa 624  
 Tyr His Tyr Val Asp Thr Lys Leu Val Ile Trp Asn Asn Asp Lys Glu  
                   195                  200                  205

ttc atg aaa gtt gag gag cat gaa atc gcc gtt gca cgc cac cat ccg 672  
 Phe Met Lys Val Glu Glu His Glu Ile Ala Val Ala Arg His His Pro  
                   210                  215                  220

ttc tat gag cca aag aag gat aag taa 699  
 Phe Tyr Glu Pro Lys Lys Asp Lys  
                   225                  230

<210> 20

<211> 232

<212> PRT

<213> Discosoma striata

<400> 20

Met Ser Cys Ser Lys Ser Val Ile Lys Glu Glu Met Leu Ile Asp Leu  
 1                  5                  10                  15

His Leu Glu Gly Thr Phe Asn Gly His Tyr Phe Glu Ile Lys Gly Lys  
                   20                  25                  30

Gly Lys Gly Gln Pro Asn Glu Gly Thr Asn Thr Val Thr Leu Glu Val  
                   35                  40                  45

Thr Lys Gly Gly Pro Leu Pro Phe Gly Trp His Ile Leu Cys Pro Gln  
                   50                  55                  60

Phe Gln Tyr Gly Asn Lys Ala Phe Val His His Pro Asp Asn Ile His  
 65                  70                  75                  80

Asp Tyr Leu Lys Leu Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ser  
                   85                  90                  95

Met His Phe Glu Asp Gly Gly Leu Cys Cys Ile Thr Asn Asp Ile Ser  
                   100                  105                  110

Leu Thr Gly Asn Cys Phe Tyr Tyr Asp Ile Lys Phe Thr Gly Leu Asn  
                   115                  120                  125

Phe Pro Pro Asn Gly Pro Val Val Gln Lys Lys Thr Thr Gly Trp Glu  
                   130                  135                  140

Pro Ser Thr Glu Arg Leu Tyr Pro Arg Asp Gly Val Leu Ile Gly Asp  
 145                  150                  155                  160

24/102

Ile His His Ala Leu Thr Val Glu Gly Gly Gly His Tyr Ala Cys Asp  
165 170 175  
Ile Lys Thr Val Tyr Arg Ala Lys Lys Ala Ala Leu Lys Met Pro Gly  
180 185 190  
Tyr His Tyr Val Asp Thr Lys Leu Val Ile Trp Asn Asn Asp Lys Glu  
195 200 205  
Phe Met Lys Val Glu Glu His Glu Ile Ala Val Ala Arg His His Pro  
210 215 220  
Phe Tyr Glu Pro Lys Lys Asp Lys  
225 230

<210> 21  
<211> 678  
<212> DNA  
<213> Discosoma sp.

<220>  
<221> CDS  
<222> (1)..(678)  
<223> fluorescent protein

<400> 21  
atg agg tct tcc aag aat gtt atc aag gag ttc atg agg ttt aag gtt 48  
Met Arg Ser Ser Lys Asn Val Ile Lys Glu Phe Met Arg Phe Lys Val  
1 5 10 15  
cgc atg gaa gga acg gtc aat ggg cac gag ttt gaa ata gaa ggc gaa 96  
Arg Met Glu Gly Thr Val Asn Gly His Glu Phe Glu Ile Glu Gly Glu  
20 25 30  
gga gag ggg agg cca tac gaa ggc cac aat acc gta aag ctt aag gta 144  
Gly Glu Gly Arg Pro Tyr Glu Gly His Asn Thr Val Lys Leu Lys Val  
35 40 45  
acc aag ggg gga cct ttg cca ttt gct tgg gat att ttg tca cca caa 192  
Thr Lys Gly Gly Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln  
50 55 60  
ttt cag tat gga agc aag gta tat gtc aag cac cct gcc gac ata cca 240  
Phe Gln Tyr Gly Ser Lys Val Tyr Val Lys His Pro Ala Asp Ile Pro  
65 70 75 80  
gac tat aaa aag ctg tca ttt cct gaa gga ttt aaa tgg gaa agg gtc 288  
Asp Tyr Lys Lys Leu Ser Phe Pro Glu Gly Phe Lys Trp Glu Arg Val  
85 90 95  
atg aac ttt gaa gac ggt ggc gtc gtt act gta acc cag gat tcc agt 336  
Met Asn Phe Glu Asp Gly Gly Val Val Thr Val Thr Gln Asp Ser Ser  
100 105 110  
ttg cag gat ggc tgt ttc atc tac aag gtc aag ttc att ggc gtg aac 384  
Leu Gln Asp Gly Cys Phe Ile Tyr Lys Val Lys Phe Ile Gly Val Asn  
115 120 125



ttt cct tcc gat gga cct gtt atg caa aag aag aca atg ggc tgg gaa	432
Phe Pro Ser Asp Gly Pro Val Met Gln Lys Lys Thr Met Gly Trp Glu	
130 135 140	
gcc agc act gag cgt ttg tat cct cgt gat ggc gtg ttg aaa gga gag	480
Ala Ser Thr Glu Arg Leu Tyr Pro Arg Asp Gly Val Leu Lys Gly Glu	
145 150 155 160	
att cat aag gct ctg aag ctg aaa gac ggt ggt cat tac cta gtt gaa	528
Ile His Lys Ala Leu Lys Leu Lys Asp Gly Gly His Tyr Leu Val Glu	
165 170 175	
ttc aaa agt att tac atg gca aag aag cct gtg cag cta cca ggg tac	576
Phe Lys Ser Ile Tyr Met Ala Lys Lys Pro Val Gln Leu Pro Gly Tyr	
180 185 190	
tac tat gtt gac tcc aaa ctg gat ata aca agc cac aac gaa gac tat	624
Tyr Tyr Val Asp Ser Lys Leu Asp Ile Thr Ser His Asn Glu Asp Tyr	
195 200 205	
aca atc gtt gag cag tat gaa aga acc gag gga cgc cac cat ctg ttc	672
Thr Ile Val Glu Gln Tyr Glu Arg Thr Glu Gly Arg His His Leu Phe	
210 215 220	
ctt taa	678
Leu	
225	

<210> 22  
 <211> 225  
 <212> PRT  
 <213> Discosoma sp.

<400> 22	
Met Arg Ser Ser Lys Asn Val Ile Lys Glu Phe Met Arg Phe Lys Val	
1 5 10 15	
Arg Met Glu Gly Thr Val Asn Gly His Glu Phe Glu Ile Glu Gly Glu	
20 25 30	
Gly Glu Gly Arg Pro Tyr Glu Gly His Asn Thr Val Lys Leu Lys Val	
35 40 45	
Thr Lys Gly Gly Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln	
50 55 60	
Phe Gln Tyr Gly Ser Lys Val Tyr Val Lys His Pro Ala Asp Ile Pro	
65 70 75 80	
Asp Tyr Lys Lys Leu Ser Phe Pro Glu Gly Phe Lys Trp Glu Arg Val	
85 90 95	
Met Asn Phe Glu Asp Gly Gly Val Val Thr Val Thr Gln Asp Ser Ser	
100 105 110	

26/102

Leu Gln Asp Gly Cys Phe Ile Tyr Lys Val Lys Phe Ile Gly Val Asn  
 115 120 125  
 Phe Pro Ser Asp Gly Pro Val Met Gln Lys Lys Thr Met Gly Trp Glu  
 130 135 140  
 Ala Ser Thr Glu Arg Leu Tyr Pro Arg Asp Gly Val Leu Lys Gly Glu  
 145 150 155 160  
 Ile His Lys Ala Leu Lys Leu Lys Asp Gly Gly His Tyr Leu Val Glu  
 165 170 175  
 Phe Lys Ser Ile Tyr Met Ala Lys Lys Pro Val Gln Leu Pro Gly Tyr  
 180 185 190  
 Tyr Tyr Val Asp Ser Lys Leu Asp Ile Thr Ser His Asn Glu Asp Tyr  
 195 200 205  
 Thr Ile Val Glu Gln Tyr Glu Arg Thr Glu Gly Arg His His Leu Phe  
 210 215 220  
 Leu  
 225

<210> 23  
 <211> 801  
 <212> DNA  
 <213> Clavularia sp.

<220>  
 <221> CDS  
 <222> (1)..(801)  
 <223> fluorescent protein

<400> 23  
 atg aag tgt aaa ttt gtg ttc tgc ctg tcc ttc ttg gtc ctc gcc atc 48  
 Met Lys Cys Lys Phe Val Phe Cys Leu Ser Phe Leu Val Leu Ala Ile  
 1 5 10 15  
 aca aac gcg aac att ttt ttg aga aac gag gct gac tta gaa gag aag 96  
 Thr Asn Ala Asn Ile Phe Leu Arg Asn Glu Ala Asp Leu Glu Glu Lys  
 20 25 30  
 aca ttg aga ata cca aaa gct cta acc acc atg ggt gtg att aaa cca 144  
 Thr Leu Arg Ile Pro Lys Ala Leu Thr Thr Met Gly Val Ile Lys Pro  
 35 40 45  
 gac atg aag att aag ctg aag atg gaa gga aat gta aac ggg cat gct 192  
 Asp Met Lys Ile Lys Leu Lys Met Glu Gly Asn Val Asn Gly His Ala  
 50 55 60  
 ttt gtg atc gaa gga gaa gga gaa gga aag cct tac gat ggg aca cac 240  
 Phe Val Ile Glu Gly Glu Gly Glu Gly Lys Pro Tyr Asp Gly Thr His  
 65 70 75 80

27/102

act tta aac ctg gaa gtg aag gaa ggt gcg cct ctg cct ttt tct tac	288
Thr Leu Asn Leu Glu Val Lys Glu Gly Ala Pro Leu Pro Phe Ser Tyr	
85 90 95	
gat atc ttg tca aac gcg ttc cag tac gga aac aga gca ttg aca aaa	336
Asp Ile Leu Ser Asn Ala Phe Gln Tyr Gly Asn Arg Ala Leu Thr Lys	
100 105 110	
tac cca gac gat ata gca gac tat ttc aag cag tcg ttt ccc gag gga	384
Tyr Pro Asp Asp Ile Ala Asp Tyr Phe Lys Gln Ser Phe Pro Glu Gly	
115 120 125	
tat tcc tgg gaa aga acc atg act ttt gaa gac aaa ggc att gtc aaa	432
Tyr Ser Trp Glu Arg Thr Met Thr Phe Glu Asp Lys Gly Ile Val Lys	
130 135 140	
gtg aaa agt gac ata agc atg gag gaa gac tcc ttt atc tat gaa att	480
Val Lys Ser Asp Ile Ser Met Glu Glu Asp Ser Phe Ile Tyr Glu Ile	
145 150 155 160	
cgt ttt gat ggg atg aac ttt cct ccc aat ggt ccg gtt atg cag aaa	528
Arg Phe Asp Gly Met Asn Phe Pro Pro Asn Gly Pro Val Met Gln Lys	
165 170 175	
aaa act ttg aag tgg gaa cca tcc act gag att atg tac gtg cgt gat	576
Lys Thr Leu Lys Trp Glu Pro Ser Thr Glu Ile Met Tyr Val Arg Asp	
180 185 190	
gga gtg ctg gtc gga gat att agc cat tct ctg ttg ctg gag gga ggt	624
Gly Val Leu Val Gly Asp Ile Ser His Ser Leu Leu Leu Glu Gly Gly	
195 200 205	
ggc cat tac cga tgt gac ttc aaa agt att tac aaa gca aaa aaa gtt	672
Gly His Tyr Arg Cys Asp Phe Lys Ser Ile Tyr Lys Ala Lys Lys Val	
210 215 220	
gtc aaa ttg cca gac tat cac ttt gtg gac cat cgc att gag atc ttg	720
Val Lys Leu Pro Asp Tyr His Phe Val Asp His Arg Ile Glu Ile Leu	
225 230 235 240	
aac cat gac aag gat tac aac aaa gta acg ctg tat gag aat gca gtt	768
Asn His Asp Lys Asp Tyr Asn Lys Val Thr Leu Tyr Glu Asn Ala Val	
245 250 255	
gct cgc tat tct ttg ctg cca agt cag gcc tag	801
Ala Arg Tyr Ser Leu Leu Pro Ser Gln Ala	
260 265	

<210> 24

<211> 266

<212> PRT

<213> Clavularia sp.

<400> 24

Met Lys Cys Lys Phe Val Phe Cys Leu Ser Phe Leu Val Leu Ala Ile
1 5 10 15

28/102

Thr	Asn	Ala	Asn	Ile	Phe	Leu	Arg	Asn	Glu	Ala	Asp	Leu	Glu	Glu	Lys
			20					25					30		
Thr	Leu	Arg	Ile	Pro	Lys	Ala	Leu	Thr	Thr	Met	Gly	Val	Ile	Lys	Pro
		35					40					45			
Asp	Met	Lys	Ile	Lys	Leu	Lys	Met	Glu	Gly	Asn	Val	Asn	Gly	His	Ala
	50					55					60				
Phe	Val	Ile	Glu	Gly	Glu	Gly	Glu	Gly	Lys	Pro	Tyr	Asp	Gly	Thr	His
65					70					75					80
Thr	Leu	Asn	Leu	Glu	Val	Lys	Glu	Gly	Ala	Pro	Leu	Pro	Phe	Ser	Tyr
			85						90					95	
Asp	Ile	Leu	Ser	Asn	Ala	Phe	Gln	Tyr	Gly	Asn	Arg	Ala	Leu	Thr	Lys
			100					105					110		
Tyr	Pro	Asp	Asp	Ile	Ala	Asp	Tyr	Phe	Lys	Gln	Ser	Phe	Pro	Glu	Gly
		115					120					125			
Tyr	Ser	Trp	Glu	Arg	Thr	Met	Thr	Phe	Glu	Asp	Lys	Gly	Ile	Val	Lys
	130					135					140				
Val	Lys	Ser	Asp	Ile	Ser	Met	Glu	Glu	Asp	Ser	Phe	Ile	Tyr	Glu	Ile
145					150					155					160
Arg	Phe	Asp	Gly	Met	Asn	Phe	Pro	Pro	Asn	Gly	Pro	Val	Met	Gln	Lys
			165						170					175	
Lys	Thr	Leu	Lys	Trp	Glu	Pro	Ser	Thr	Glu	Ile	Met	Tyr	Val	Arg	Asp
		180						185					190		
Gly	Val	Leu	Val	Gly	Asp	Ile	Ser	His	Ser	Leu	Leu	Leu	Glu	Gly	Gly
		195					200					205			
Gly	His	Tyr	Arg	Cys	Asp	Phe	Lys	Ser	Ile	Tyr	Lys	Ala	Lys	Lys	Val
	210					215					220				
Val	Lys	Leu	Pro	Asp	Tyr	His	Phe	Val	Asp	His	Arg	Ile	Glu	Ile	Leu
225					230					235					240
Asn	His	Asp	Lys	Asp	Tyr	Asn	Lys	Val	Thr	Leu	Tyr	Glu	Asn	Ala	Val
			245					250						255	
Ala	Arg	Tyr	Ser	Leu	Leu	Pro	Ser	Gln	Ala						
			260					265							

<210> 25

<211> 881

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 25

Met	Lys	Leu	Leu	Ser	Ser	Ile	Glu	Gln	Ala	Cys	Asp	Ile	Cys	Arg	Leu
1				5					10					15	

Lys	Lys	Leu	Lys 20	Cys	Ser	Lys	Glu	Lys 25	Pro	Lys	Cys	Ala	Lys 30	Cys	Leu
Lys	Asn	Asn 35	Trp	Glu	Cys	Arg	Tyr 40	Ser	Pro	Lys	Thr	Lys 45	Arg	Ser	Pro
Leu	Thr 50	Arg	Ala	His	Leu	Thr 55	Glu	Val	Glu	Ser	Arg 60	Leu	Glu	Arg	Leu
Glu 65	Gln	Leu	Phe	Leu 70	Leu	Ile	Phe	Pro	Arg	Glu 75	Asp	Leu	Asp	Met	Ile 80
Leu	Lys	Met	Asp	Ser 85	Leu	Gln	Asp	Ile	Lys 90	Ala	Leu	Leu	Thr	Gly 95	Leu
Phe	Val	Gln	Asp 100	Asn	Val	Asn	Lys 105	Asp	Ala	Val	Thr	Asp 110	Arg	Leu	Ala
Ser	Val	Glu 115	Thr	Asp	Met	Pro	Leu 120	Thr	Leu	Arg	Gln	His 125	Arg	Ile	Ser
Ala	Thr 130	Ser	Ser	Ser	Glu	Glu 135	Ser	Ser	Asn	Lys	Gly 140	Gln	Arg	Gln	Leu
Thr 145	Val	Ser	Ile	Asp 150	Ser	Ala	Ala	His	His	Asp 155	Asn	Ser	Thr	Ile	Pro 160
Leu	Asp	Phe	Met	Pro 165	Arg	Asp	Ala	Leu	His 170	Gly	Phe	Asp	Trp	Ser 175	Glu
Glu	Asp	Asp	Met 180	Ser	Asp	Gly	Leu	Pro 185	Phe	Leu	Lys	Thr	Asp 190	Pro	Asn
Asn	Asn	Gly 195	Phe	Phe	Gly	Asp	Gly 200	Ser	Leu	Leu	Cys	Ile 205	Leu	Arg	Ser
Ile	Gly 210	Phe	Lys	Pro	Glu	Asn 215	Tyr	Thr	Asn	Ser	Asn 220	Val	Asn	Arg	Leu
Pro 225	Thr	Met	Ile	Thr 230	Asp	Arg	Tyr	Thr	Leu	Ala 235	Ser	Arg	Ser	Thr	Thr 240
Ser	Arg	Leu	Leu	Gln 245	Ser	Tyr	Leu	Asn	Asn 250	Phe	His	Pro	Tyr	Cys 255	Pro
Ile	Val	His	Ser 260	Pro	Thr	Leu	Met	Met 265	Leu	Tyr	Asn	Asn	Gln 270	Ile	Glu
Ile	Ala	Ser 275	Lys	Asp	Gln	Trp	Gln 280	Ile	Leu	Phe	Asn	Cys 285	Ile	Leu	Ala
Ile	Gly 290	Ala	Trp	Cys	Ile	Glu 295	Gly	Glu	Ser	Thr	Asp 300	Ile	Asp	Val	Phe
Tyr 305	Tyr	Gln	Asn	Ala 310	Lys	Ser	His	Leu	Thr	Ser 315	Lys	Val	Phe	Glu	Ser 320

Gly	Ser	Ile	Ile	Leu	Val	Thr	Ala	Leu	His	Leu	Leu	Ser	Arg	Tyr	Thr	325	330	335
Gln	Trp	Arg	Gln	Lys	Thr	Asn	Thr	Ser	Tyr	Asn	Phe	His	Ser	Phe	Ser	340	345	350
Ile	Arg	Met	Ala	Ile	Ser	Leu	Gly	Leu	Asn	Arg	Asp	Leu	Pro	Ser	Ser	355	360	365
Phe	Ser	Asp	Ser	Ser	Ile	Leu	Glu	Gln	Arg	Arg	Arg	Ile	Trp	Trp	Ser	370	375	380
Val	Tyr	Ser	Trp	Glu	Ile	Gln	Leu	Ser	Leu	Leu	Tyr	Gly	Arg	Ser	Ile	385	390	395
Gln	Leu	Ser	Gln	Asn	Thr	Ile	Ser	Phe	Pro	Ser	Ser	Val	Asp	Asp	Val	405	410	415
Gln	Arg	Thr	Thr	Thr	Gly	Pro	Thr	Ile	Tyr	His	Gly	Ile	Ile	Glu	Thr	420	425	430
Ala	Arg	Leu	Leu	Gln	Val	Phe	Thr	Lys	Ile	Tyr	Glu	Leu	Asp	Lys	Thr	435	440	445
Val	Thr	Ala	Glu	Lys	Ser	Pro	Ile	Cys	Ala	Lys	Lys	Cys	Leu	Met	Ile	450	455	460
Cys	Asn	Glu	Ile	Glu	Glu	Val	Ser	Arg	Gln	Ala	Pro	Lys	Phe	Leu	Gln	465	470	475
Met	Asp	Ile	Ser	Thr	Thr	Ala	Leu	Thr	Asn	Leu	Leu	Lys	Glu	His	Pro	485	490	495
Trp	Leu	Ser	Phe	Thr	Arg	Phe	Glu	Leu	Lys	Trp	Lys	Gln	Leu	Ser	Leu	500	505	510
Ile	Ile	Tyr	Val	Leu	Arg	Asp	Phe	Phe	Thr	Asn	Phe	Thr	Gln	Lys	Lys	515	520	525
Ser	Gln	Leu	Glu	Gln	Asp	Gln	Asn	Asp	His	Gln	Ser	Tyr	Glu	Val	Lys	530	535	540
Arg	Cys	Ser	Ile	Met	Leu	Ser	Asp	Ala	Ala	Gln	Arg	Thr	Val	Met	Ser	545	550	555
Val	Ser	Ser	Tyr	Met	Asp	Asn	His	Asn	Val	Thr	Pro	Tyr	Phe	Ala	Trp	565	570	575
Asn	Cys	Ser	Tyr	Tyr	Leu	Phe	Asn	Ala	Val	Leu	Val	Pro	Ile	Lys	Thr	580	585	590
Leu	Leu	Ser	Asn	Ser	Lys	Ser	Asn	Ala	Glu	Asn	Asn	Glu	Thr	Ala	Gln	595	600	605
Leu	Leu	Gln	Gln	Ile	Asn	Thr	Val	Leu	Met	Leu	Leu	Lys	Lys	Leu	Ala	610	615	620

Thr Phe Lys Ile Gln Thr Cys Glu Lys Tyr Ile Gln Val Leu Glu Glu  
 625 630 635 640  
 Val Cys Ala Pro Phe Leu Leu Ser Gln Cys Ala Ile Pro Leu Pro His  
 645 650 655  
 Ile Ser Tyr Asn Asn Ser Asn Gly Ser Ala Ile Lys Asn Ile Val Gly  
 660 665 670  
 Ser Ala Thr Ile Ala Gln Tyr Pro Thr Leu Pro Glu Glu Asn Val Asn  
 675 680 685  
 Asn Ile Ser Val Lys Tyr Val Ser Pro Gly Ser Val Gly Pro Ser Pro  
 690 695 700  
 Val Pro Leu Lys Ser Gly Ala Ser Phe Ser Asp Leu Val Lys Leu Leu  
 705 710 715 720  
 Ser Asn Arg Pro Pro Ser Arg Asn Ser Pro Val Thr Ile Pro Arg Ser  
 725 730 735  
 Thr Pro Ser His Arg Ser Val Thr Pro Phe Leu Gly Gln Gln Gln Gln  
 740 745 750  
 Leu Gln Ser Leu Val Pro Leu Thr Pro Ser Ala Leu Phe Gly Gly Ala  
 755 760 765  
 Asn Phe Asn Gln Ser Gly Asn Ile Ala Asp Ser Ser Leu Ser Phe Thr  
 770 775 780  
 Phe Thr Asn Ser Ser Asn Gly Pro Asn Leu Ile Thr Thr Gln Thr Asn  
 785 790 795 800  
 Ser Gln Ala Leu Ser Gln Pro Ile Ala Ser Ser Asn Val His Asp Asn  
 805 810 815  
 Phe Met Asn Asn Glu Ile Thr Ala Ser Lys Ile Asp Asp Gly Asn Asn  
 820 825 830  
 Ser Lys Pro Leu Ser Pro Gly Trp Thr Asp Gln Thr Ala Tyr Asn Ala  
 835 840 845  
 Phe Gly Ile Thr Thr Gly Met Phe Asn Thr Thr Thr Met Asp Asp Val  
 850 855 860  
 Tyr Asn Tyr Leu Phe Asp Asp Glu Asp Thr Pro Pro Asn Pro Lys Lys  
 865 870 875 880

Glu

<210> 26

<211> 202

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 26  
 Met Lys Ala Leu Thr Ala Arg Gln Gln Glu Val Phe Asp Leu Ile Arg  
 1 5 10 15  
 Asp His Ile Ser Gln Thr Gly Met Pro Pro Thr Arg Ala Glu Ile Ala  
 20 25 30  
 Gln Arg Leu Gly Phe Arg Ser Pro Asn Ala Ala Glu Glu His Leu Lys  
 35 40 45  
 Ala Leu Ala Arg Lys Gly Val Ile Glu Ile Val Ser Gly Ala Ser Arg  
 50 55 60  
 Gly Ile Arg Leu Leu Gln Glu Glu Glu Gly Leu Pro Leu Val Gly  
 65 70 75 80  
 Arg Val Ala Ala Gly Glu Pro Leu Leu Ala Gln Gln His Ile Glu Gly  
 85 90 95  
 His Tyr Gln Val Asp Pro Ser Leu Phe Lys Pro Asn Ala Asp Phe Leu  
 100 105 110  
 Leu Arg Val Ser Gly Met Ser Met Lys Asp Ile Gly Ile Met Asp Gly  
 115 120 125  
 Asp Leu Leu Ala Val His Lys Thr Gln Asp Val Arg Asn Gly Gln Val  
 130 135 140  
 Val Val Ala Arg Ile Asp Asp Glu Val Thr Val Lys Arg Leu Lys Lys  
 145 150 155 160  
 Gln Gly Asn Lys Val Glu Leu Leu Pro Glu Asn Ser Glu Phe Lys Pro  
 165 170 175  
 Ile Val Val Asp Leu Arg Gln Gln Ser Phe Thr Ile Glu Gly Leu Ala  
 180 185 190  
 Val Gly Val Ile Arg Asn Gly Asp Trp Leu  
 195 200

<210> 27

<211> 17

<212> DNA

<213> Unknown Organism

<220>

<223> Description of Unknown Organism: Nucleotide sequence that  
 is bound by GAL4

<220>

<221> modified\_base

<222> (4)..(4)

<223> any purine



<220>  
 <221> modified\_base  
 <222> (5)..(5)  
 <223> any nucleotide

<220>  
 <221> modified\_base  
 <222> (6)..(6)  
 <223> any nucleotide

<220>  
 <221> modified\_base  
 <222> (7)..(7)  
 <223> any purine

<220>  
 <221> modified\_base  
 <222> (9)..(9)  
 <223> any pyrimidine

<220>  
 <221> modified\_base  
 <222> (10)..(10)  
 <223> any nucleotide

<220>  
 <221> modified\_base  
 <222> (11)..(11)  
 <223> any pyrimidine

<220>  
 <221> modified\_base  
 <222> (12)..(12)  
 <223> any nucleotide

<220>  
 <221> modified\_base  
 <222> (14)..(14)  
 <223> any nucleotide

<400> 27  
 cggnnnnncnn nncnccg

17

<210> 28  
 <211> 14  
 <212> DNA  
 <213> Unknown Organism

<220>  
 <223> Description of Unknown Organism: Nucleotide sequence  
 that binds LexA

<220>  
 <221> modified\_base  
 <222> (6)..(9)  
 <223> any nucleotide

<400> 28  
cgaacnnnng ttcg 14

<210> 29  
<211> 25  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic primer

<400> 29  
ctagacgttg acattgatta ttgac 25

<210> 30  
<211> 25  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic primer

<400> 30  
tctagagcca gtaagcagtg ggttc 25

<210> 31  
<211> 28  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic primer

<400> 31  
gcggccgccc attgcattgg gattggtc 28

<210> 32  
<211> 26  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic primer

<400> 32  
gcggccgccc tccaagtctt tatctg 26

<210> 33  
<211> 28  
<212> DNA  
<213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: Synthetic primer  
  
 <400> 33  
 ggtaccggct cgtagagaaa tatgaacc 28  
  
 <210> 34  
 <211> 26  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> Description of Artificial Sequence: Synthetic primer  
  
 <400> 34  
 ctcgagagag actgaatttc cctttt 26  
  
 <210> 35  
 <211> 30  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> Description of Artificial Sequence: Synthetic primer  
  
 <400> 35  
 gcggccgcga aggaagatag ccaaggctta 30  
  
 <210> 36  
 <211> 32  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> Description of Artificial Sequence: Synthetic primer  
  
 <400> 36  
 gcggccgctt aaagatgcca cactgctggt ct 32  
  
 <210> 37  
 <211> 26  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> Description of Artificial Sequence: Synthetic primer  
  
 <400> 37  
 gtcgacacct gtcggctact gggata 26  
  
 <210> 38  
 <211> 26  
 <212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic primer

<400> 38

gatatcgtgg gactctgttt ccagca

26

<210> 39

<211> 6795

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic recombinant DNA vector sequence

<220>

<221> modified\_base

<222> (892)..(892)

<223> unknown nucleotide

<400> 39

aagtc	cg	gca	cctc	gtgc	ac	gcgg	atttc	g	gtcc	caaca	a	tg	tc	ctg	ac	g	aca	atg	ggcc	60			
gcata	ac	agc	g	ggtc	attg	ac	tgga	gcg	agg	cgat	gttc	gg	attc	ccca	tac	gag	gtc		120				
ccaac	atc	ctt	ctt	ctg	gag	g	ccgt	ggt	tg	ctt	gtat	gga	gcag	cag	ac	g	ctact	tc	180				
agc	g	gag	gca	tcc	gag	ctt	gcag	gat	cgc	cgc	ggct	ccg	ggc	gtat	atg	ctcc	gcatt	g	240				
gtc	ttg	acca	act	ctat	cag	ag	ctt	ggt	tg	ac	gg	caatt	cgat	gat	gca	gctt	gggc	gc	300				
aggg	tc	gatg	cg	ac	gca	atc	gtcc	gat	ccg	gag	ccggg	ac	tg	tc	gggc	gt	acac	aaat	360				
ccc	gc	aga	ag	cgc	ggc	cg	tc	tgga	acc	gat	g	gct	gt	gtaga	agt	act	cgc	gat	agt	420			
acc	gac	gccc	cag	ca	ctc	gt	ccg	aggg	gcaa	agga	atgc	ct	gaga	aa	ggaa	gt	gag	ct	gta	480			
aagg	ct	gagc	tct	ctc	ctc	tg	ac	gtat	gtag	cct	ctg	ggt	gct	tc	gtc	ac	tact	gtt	ct	540			
tgact	ca	gca	tg	gca	atc	tg	atg	aa	tccc	ag	ctg	ta	agt	ctg	cata	aa	t	tgat	gat	cta	600		
ttaa	aca	ata	aag	at	gt	tcca	ctaa	aat	gga	ag	tttt	ttt	ac	tg	tc	ata	ctt	tg	tta	aga	660		
ggt	gaga	aaca	gag	tac	ctac	at	ttt	tga	atg	ga	agg	att	gg	ag	ctac	ggg	gt	ggg	gg	g	720		
ggg	tgg	gatt	ag	ata	aat	gc	ctg	ctc	ttta	ctg	aag	gct	c	ttt	act	att	g	ctt	t	atg	780		
atg	ttt	cata	gtt	gg	ata	tc	ata	at	tttaa	ca	ag	caaaa	ac	caa	at	ta	agg	gcc	ag	ct	cat	840	
tct	cc	actc	at	gat	ctata	gat	ctat	aga	tct	ctc	gtg	gat	cat	gtt	tnt	ct	gat	cc			900		
act	gga	agct	tat	cg	ata	cc	gtc	ga	ctc	g	aggg	ggg	g	ccg	ta	acc	ag	ctt	tt	gtt	cc	960	
ctt	tag	tgag	ggt	ta	att	g	gcg	ctt	ggc	ta	at	cat	gg	cat	ag	ct	gtt	t	ct	gt	gta	1020	
aatt	gtt	atc	cg	tc	caca	at	tcc	ac	aca	ac	ata	c	gag	ccg	ga	ag	cata	aaa	gt	g	taa	agg	1080
tggg	gt	gc	ct	aat	gag	tgag	cta	act	caca	tta	att	g	cg	t	act	g	cccc	g	ctt	t	cc	1140	
cag	tc	ggg	aa	ac	ct	gtc	gtg	cc	ag	ct	g	cat	ta	at	ga	at	cg	gcca	ac	g	cg	1200	
ggt	tt	gc	gta	tt	ggg	cg	ctc	ttc	cg	ct	c	act	g	ct	g	ct	g	ctc	gg	t	cg	1260	
cgg	ct	gc	ggc	gag	cg	gtat	c	ag	ct	cact	c	aag	gc	g	ta	a	ac	g	gtt	at	c	1320	
ggg	gata	acg	cag	gaa	agaa	cat	gtg	ag	ca	aa	agg	cc	agc	aaa	agg	cc	ag	ga	acc	g	taaaa	1380	
aagg	cc	gc	gt	tg	ct	ggc	gtt	ttt	ccat	agg	ctc	gc	cccc	ct	gac	gag	ca	tc	ac	aaaa	aat	1440	
cg	ac	gt	caa	gt	ca	gg	gtg	gc	gaa	ac	ccg	ac	agg	act	at	aa	ag	ata	cca	gg	ctt	1500	
cct	gga	ag	ct	cc	tc	gt	gcg	ctc	ct	gt	gtt	ccg	ac	cc	ct	g	cg	ctt	acc	gg	at	1560	
gc	ctt	t	ctcc	ctt	cg	gga	ag	cg	t	ggc	g	ctt	t	ct	cat	ag	ct	c	gt	at	ct	1620	
tc	gg	t	gtag	tc	gt	tc	g	ctc	ca	ag	ct	ggg	tg	gt	g	cac	a	cccc	cc	gt	1680		
cg	ct	gc	gc	ct	cc	g	gta	a	ag	ca	ac	cg	gt	ta	ag	aca	cg	act	t	at	cg	1740	
cc	act	gg	cag	cc	act	gg	ta	ac	agg	att	ag	cag	ag	cga	gg	t	at	g	ta	g	ta	1800	
gag	tt	ct	tga	agt	gg	tgg	cc	ta	act	ac	gg	c	tac	act	ag	aa	gg	ac	g	at	ct	1860	
g	ct	ct	g	ctga	ag	cc	ag	tt	ac	ctt	cg	gaaa	aa	ag	at	t	gg	t	at	ct	g	1920	
acc	ac	gc	ctg	g	tag	cg	gtg	g	ttt	ttt	gtt	tg	ca	ag	cag	c	ag	att	ac	gc	g	1980	
g	at	ct	ca	ag	a	ag	at	c	ctt	t	ct	ac	gg	g	gt	ctg	ac	g	ct	ca	gtg	2040	

tcacgttaag	ggatttttgg	catgagatta	tcaaaaagga	tcttcaccta	gaccccttta	2100
aattaaaaat	gaagttttta	atcaatctaa	agtatatatg	agtaaacttg	gtctgacagt	2160
taccaatgct	taatcagtg	ggcacctatc	tcagcgatct	gtctatttcg	ttcatccata	2220
gttgccctgac	tccccgtcgt	gtagataaact	acgatacggg	agggccttacc	atctggcccc	2280
agtgtgcaa	tgataccgcg	agacccacgc	tcaccggctc	cagatttatc	agcaataaac	2340
cagccagccg	gaagggccga	gcgagaaagt	ggctctgcaa	ctttatccgc	ctccatccag	2400
tctattaatt	gttgccggga	agctagagta	agtagttcgc	cagttaatag	tttgcgcaac	2460
gttggttgcca	ttgctacagg	catcgtgggtg	tcacgctcgt	cgtttggtat	ggcttcattc	2520
agctccgggt	cccaacgatc	aaggcgagtt	acatgatccc	ccatggttg	caaaaaagcg	2580
gttagctcct	tcggtcctcc	gatcgttggtc	agaagtaagt	tggccgcagt	gttatcactc	2640
atggttatgg	cagcactgca	taattctctt	actgtcatgc	catccgtaag	atgcttttct	2700
gtgactgggtg	agtactcaac	caagtcattc	tgagaatagt	gtatgcggcg	accgagttgc	2760
tcttgcccgg	cgtcaatacg	ggataatacc	gcgccacata	gcagaacttt	aaaagtgtct	2820
atcattggaa	aacgtttctt	ggggcgaaaa	ctctcaagga	tcttaccgct	gttgagatcc	2880
agttcgaagt	aaccactcgt	tgcacccaac	tgatcttcag	catcttttac	tttcaccagc	2940
gtttctgggt	gagcaaaaac	aggaaggcaa	aatgccgcaa	aaaagggaat	aaggcgaca	3000
cggaaatgtt	gaatactcat	actcttcctt	tttcaatatt	attgaagcat	ttatcaggggt	3060
tattgtctca	tgagcggata	catatttgaa	tgtatttaga	aaaataaaca	aataggggtt	3120
ccgcgcacat	ttccccgaaa	agtgccacct	aaattgtaag	cgtaaatatt	ttgttaaaat	3180
tcgcgttaaa	tttttggttaa	atcagctcat	tttttaacca	ataggccgaa	atcggcaaaa	3240
tcccttataa	atcaaaagaa	tagaccgaga	taggggttgag	tgttggtcca	gtttggaaca	3300
agagttcact	attaaagaac	gtggactcca	acgtcaaaag	gcgaaaaacc	gtctaatcag	3360
gcgatggccc	actacgtgaa	ccatcacctt	aatcaagttt	tttggggctg	aggtgccgta	3420
aagcactaaa	tcggaaccct	aaagggagcc	cccgatttag	agcttgacgg	ggaaagccgg	3480
cgaacgtggc	gagaaaggaa	gggaagaaag	cgaaggagc	gggcgctagg	gcgctggcaa	3540
gtgtagcgg	cacgtcgcgc	gtaaccacca	caccgcgcgc	gcttaatgcg	ccgctacagg	3600
gcgcgtccca	ttcgccattc	aggctgcgca	actggtggga	agggcgatcg	gtgcgggcct	3660
cttcgctatt	acgcgagctg	gcgaaagggg	gatgtgctgc	aaggcgatta	agttgggtaa	3720
cgccagggtt	ttcccagtca	cgagcttgta	aaactgacgc	cagtgcgcgc	gcgtaatcag	3780
actcactata	gggcgaattg	gagctccacc	gcgggtgcgc	cgggccatgc	aggccacgac	3840
atgataagat	acattgatga	gtttggacaa	accacaacta	gaatgcagtg	aaaaaaatgc	3900
tttattttgtg	aaatttgtga	tgctattgct	ttattttgtaa	ccattataag	ctgcaataaa	3960
caagttcttg	tacagctcgt	ccatgccgag	agtgatcccg	gcggcgggtca	cgaactccag	4020
caggaccatg	tgatcgcgct	tctcgttgct	gtcttcgctc	agggcggact	ggtagctcag	4080
gtagtgggtg	tcgggcagca	gcacgggggc	gtcgcgcatg	gggggtgtct	gctggtagtg	4140
gtcggcgagc	tgacgctgc	cgtcctcgat	gttggtggcg	atcttgaagt	tcaccttgat	4200
gccgttcttc	tgcttgctcg	ccatgatata	gacgttggtg	ctggttgagt	tgtactccag	4260
cttggtcccc	aggatggtgc	cgtcctcctt	gaagtcgatg	cccttcagct	cgatgcgggt	4320
caccaggggtg	tcgccctcga	acttcacctc	ggcgcggggtc	ttgtagtgtc	cgtcgtcctt	4380
gaagaagatg	gtgcgctcct	ggacgtagcc	ttcgggcatg	gcggacttga	agaagtcgtg	4440
ctgcttcatg	tggtcggggg	agcgggcgaa	gcactgcagg	ccgtagccga	aggtgtcac	4500
gaggggtgggc	cagggcacgg	gcagcttgcc	ggtggtgcag	atgaacttca	gggtcagctt	4560
gccgtaggtg	gcacgcctt	cgccctcgcc	ggacacgctg	aacttggtgg	cgtttacgtc	4620
gccgtccagc	tcgaccagga	tgggcaccac	cccggtgaac	agctcctcgc	ccttgctcac	4680
catggtggcg	ttagcttgat	ttgacagtgg	ctgggggttg	cgccgccggg	ttttatagga	4740
agccacagcg	gccactcgag	ccataaaaag	caacttttag	aacggcgggg	ggtgattgga	4800
ttcgagtctg	ttattcaccg	gccttgccgc	acagtgcagc	atttttttac	ccctctccc	4860
ctccttttgc	gggggaaaaa	aaaaaaaaaa	aaaaaaaaag	gagaagagaa	aaaaagcgag	4920
cgagagagaa	agcgagattg	aggaagagga	tgaagagttt	ggcgatgggt	gctggttccg	4980
tagggccaga	tggacaagaa	tagccccgc	ccttgcgga	agtatcccat	tcagtgactc	5040
agatcagatc	aagcggccgc	tctagaacta	gtggatccgg	ataagccagt	aagcagtggtg	5100
ttctctagtt	agccagagag	ctctgcttat	atagacctcc	caccgtacac	gcctaccgcc	5160
catttgcgtc	aatggggcgg	agttgttacg	acattttgga	aagtcctggt	gattttgggtg	5220
ccaaaacaaa	ctcccattga	cgtcaatggg	gtggagactt	ggaaatcccc	gtgagtcaaa	5280
ccgctatcca	cgcccattga	tgtactgcca	aaaccgcctc	accatggtaa	tagcgatgac	5340
taatacgtag	atgtactgcc	aagtaggaaa	gtcccataag	gtcatgtact	gggcataatg	5400
ccaggcgggc	cattttaccgt	cattgacgtc	aatagggggc	gtacttggca	tatgatacac	5460
ttgatgtact	gccaaagtgg	cagtttaccg	taaatactcc	accatttgac	gtcaatggaa	5520

agtccttatt	ggcgttacta	tgggaacata	cgctcattatt	gacgtcaatg	ggcgggggtc	5580
gttgggagg	cagccaggcg	ggccattttac	cgtaagttat	gtaacgcgga	actccatata	5640
tgggctatga	actaatgacc	ccgtaattga	ttactattaa	taactagtca	ataatcaatg	5700
tcaaccccg	gctgcaggaa	ttctaccggg	taggggaggc	gcttttccca	aggcagtctg	5760
gagcatgcgc	tttagcagcc	ccgctggcac	ttggcgcatc	acaagtggcc	tctggcctcg	5820
cacacattcc	acatccaccg	gtagcgccaa	ccggctccct	tctttggtgg	ccccttcgcg	5880
ccaccttcta	ctcctcccct	agtcagggaag	ttcccccccg	ccccgcagct	cgcgctcgtc	5940
aggacgtgac	aaatggaaat	agcacgtctc	actagtctcg	tgcagatgga	caagcaccgc	6000
tgagcaatgg	aagcgggtag	gcctttgggg	cagcggccaa	tagcagcttg	gctccttcgc	6060
tttctgggct	cagaggctgg	gaaggggtgg	gtccgggggc	gggctcaggg	gcgggctcag	6120
gggcggggcg	ggcgcaagg	tcctccggac	ccggcattct	gcacgcttca	aaagcgcacg	6180
tctgccgcgc	tgttctcctc	ttcctcatct	ccgggccttc	gacctgcatg	aaaaagcctg	6240
aactcaccgc	gacgtctgtc	gagaagtttc	tgatcgaaaa	gttcgacagc	gtctccgacc	6300
tgatgcagct	ctcggagggc	gaagaatctc	gtgcttttcag	cttcgatgta	ggagggcgctg	6360
gatatgtcct	gcgggtaaat	agctgcgcgc	atggtttcta	caaagatcgt	tatgtttatc	6420
ggcactttgc	atcggccgcg	ctcccgaatt	cggaaagtgc	tgacattggg	gaattcagcg	6480
agagcctgac	ctattgcatc	tcccgcctg	cacagggtgt	cacgttgcaa	gacctgcctg	6540
aaaccgaact	gcccgcgtgt	ctgcagccgc	tgcggagggc	catggatgcg	atcgctgcgg	6600
ccgatcttag	ccagacgagc	gggttcggcc	cattcggacc	gcaaggaatc	ggtcaataca	6660
ctacatggcg	tgatttcata	tgcgcgattg	ctgatcccca	tgtgtatcac	tggtcaactg	6720
tgatggacga	caccgtcagt	gcgtccgctc	cgcaggctct	cgatgagctg	atgctttggg	6780
ccgaggactg	ccccg					6795

&lt;210&gt; 40

&lt;211&gt; 5821

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence: Synthetic recombinant  
DNA vector sequence

&lt;400&gt; 40

ggatctcaag	aagatccttt	gatcttttct	acgggggtctg	acgctcagtg	gaacgaaaac	60
tcacgttaag	ggatttttgt	catgagatta	tcaaaaagga	tcttcaccta	gatcctttta	120
aattaaaaat	gaagttttta	atcaatctaa	agtatatatg	agtaaaactg	gtctgacagt	180
taccaatgct	taatcagtga	ggcacctatc	tcagcgatct	gtctatttcg	ttcatccata	240
gttgccctgac	tccccgctcg	gtagataact	acgatacggg	agggccttacc	atctggcccc	300
agtgtgcaa	tgataccgcg	agaccacgc	tcaccggctc	cagatttatc	agcaataaac	360
cagccagccg	gaagggccga	gcgcagaagt	ggctctgcaa	ctttatccgc	ctccatccag	420
tctattaatt	gttgccggga	agctagagta	agtagttcgc	cagttaatag	tttgcgcaac	480
gttggtgcca	ttgctacagg	catcgtgggt	tcacgctcgt	cgtttggtat	ggcttcattc	540
agctccggtt	cccaacgatc	aaggcgagtt	acatgatccc	ccatgttgtg	caaaaaagcg	600
gttagctcct	tcggctcctc	gatcgttgtc	agaagtaagt	tggccgcagt	gttatcactc	660
atggttatgg	cagcactgca	taattctctt	actgtcatgc	catccgtaag	atgcttttct	720
gtgactgggt	agtactcaac	caagtcattc	tgagaatagt	gtatgcggcg	accgagttgc	780
tcttgcccgg	cgtcaatacg	ggataatacc	gcgccacata	gcagaacttt	aaaagtgtct	840
atcattggaa	aacgtttctt	ggggcgaaaa	ctctcaagga	tcttaccgct	gttgagatcc	900
agttcgatgt	aaccctactc	tgcacccaac	tgatcttcag	catcttttac	tttcaccagc	960
gtttctgggt	gagcaaaaaa	aggaaggcaa	aatgccgcaa	aaaaggggaat	aagggcgaca	1020
cggaaatgtt	gaatactcat	actcttctct	tttcaatatt	attgaagcat	ttatcagggt	1080
tattgtctca	tgagcggata	catatttgaa	tgtatttaga	aaaataaaca	aataggggtt	1140
ccgcgcacat	ttccccgaaa	agtgccacct	gacgtcgacg	gatcgggaga	tctcccgatc	1200
ccctatgggt	cactctcagt	acaatctgct	ctgatgccgc	atagttaagc	cagtatctgc	1260
tccctgcttg	tgtgttggtg	gtcgctgagt	agtgcgcgag	caaaatttaa	gctacaacaa	1320
ggcaaggcct	gaccgacaat	tgcatagaag	atctgcttag	ggtaggcgt	tttgcgctgc	1380
ttcgcgatgt	acggggccaga	tatacgcggt	gacattgatt	attgactagt	tattaatagt	1440

aatcaattac	ggggtcatta	gttcatagcc	catatatgga	gttccgcgtt	acataactta	1500
cggtaaattg	cccgcctggc	tgaccgcccc	acgacccccg	cccattgacg	tcaataatga	1560
cgtatgttcc	catagtaacg	ccaatagggg	ctttccattg	acgtcaatgg	gtggagtatt	1620
tacggtaaac	tgcccacttg	gcagtacatc	aagtgtatca	tatgccaaagt	acgcccccta	1680
ttgacgtcaa	tgacggtaaa	tggcccgccct	ggcattatgc	ccagtacatg	accttatggg	1740
acttttctac	ttggcagtac	atctacgtat	tagtcatcgc	tattaccatg	gtgatgcggt	1800
tttggcagta	catcaatggg	cgtggatagc	ggtttgactc	acggggattt	ccaagtctcc	1860
accccattga	cgtcaatggg	agtttgtttt	ggcaccaaaa	tcaacgggac	tttccaaaat	1920
gtcgtaaaca	ctccgcccc	ttgacgcaaa	tgggcggtag	gcgtgtacgg	tgggaggtct	1980
atataagcag	agctctctgg	ctaactagag	aaccactgc	ttactggctt	atcgaaatta	2040
atacgactca	ctatagggag	acccaagctg	gctagcgttt	aaacttaagc	ttgccaccat	2100
gaagctactg	tcttctatcg	aacaagcatg	cgatatttgc	cgacttaaaa	agctcaagtg	2160
ctccaaagaa	aaaccgaagt	gcgccaagtg	tctgaagaac	aactgggagt	gtcgtactct	2220
tcccaaaacc	aaaagggtct	cgtgtgactag	ggcacatctg	acagaagtgg	aatcaaggct	2280
agaaagactg	gaacagctat	ttctactgat	ttttcctcga	gaagacctg	acatgatttt	2340
gaaaatggat	tctttacagg	atataaaagc	attgttaaca	ggattatttg	tacaagataa	2400
tgtgaataaa	gatgccgtca	cagatagatt	ggcttcagtg	gagactgata	tgctctaac	2460
attgagacag	ttgactgtat	cgccgggtac	cgagctcgga	tccactagtc	cagtgtgggtg	2520
gaattctgca	gatatccagc	acagtggcgg	ccgctcgagt	ctagagggcc	cgtttaaacc	2580
cgtgatcag	cctcgactgt	gccttctagt	tgccagccat	ctgttgtttg	cccccccc	2640
gtgccttctt	tgaccctgga	aggtgccact	cccactgtcc	tttctctaata	aaatgaggaa	2700
attgcactcg	attgtctgag	taggtgtcat	gttattctgg	ggggtggggg	ggggcaggac	2760
agcaaggggg	aggattggga	agacaatagc	aggcatgctg	gggatgcggt	gggctctatg	2820
gcttctgagg	cggaaagaac	cagctggggc	tctagggggg	atccccacgc	gccctgtagc	2880
ggcgcattaa	gcgcggcggg	tgtgggtggt	acgcgcagcg	tgaccgctac	acttgccagc	2940
gccctagcgc	ccgctccttt	cgttttcttc	ccttcctttc	tgcacacgtt	cgccggcttt	3000
ccccgtcaag	ctctaaatcg	ggggctccct	ttagggttcc	gatttagtgc	tttacggcac	3060
ctcgacccca	aaaaacttga	ttagggtagt	ggttcacgta	gtggggccatc	gccctgatag	3120
acggtttttt	caactttgac	gttggagtcc	acgttcttta	atagtggact	cttgttccaa	3180
actggaacaa	cactcaaccc	tatctcggtc	tattcttttg	atttataagg	gattttgccg	3240
atttcggcct	attggttaaa	aaatgagctg	atttaacaaa	aatttaacgc	gaattaattc	3300
tgtggaatgt	gtgtcagtta	gggtgtggaa	agtccccagg	ctccccagca	ggcagaagta	3360
tgcaaagcat	gcatctcaat	tagtcagcaa	ccagggtgtg	aaagtcccca	ggctccccag	3420
caggcagaag	tatgcaaagc	atgcactcca	attagtcagc	aaccatagtc	ccgcccctaa	3480
ctccgcccac	cccgcgccca	actccgcccc	gttccgcccc	ttctccgccc	catggctgac	3540
taattttttt	tattttatgca	gaggccgagg	ccgcctctgc	ctctgagcta	ttccagaagt	3600
agtgaggagg	cttttttgga	ggcctaggct	tttgcaaaaa	gctcccggga	gcttgatat	3660
ccattttctg	atctgatcaa	gagacaggat	gaggatcggt	tgcgatgatt	gaacaagatg	3720
gattgcacgc	aggttctccg	gccgcttggg	tggagaggct	attcggctat	gactgggcac	3780
aacagacaat	cggctgctct	gatgccgcgc	tgttccggct	gtcagcgcag	gggcgcccgg	3840
ttctttttgt	caagaccgac	ctgtccgggt	ccctgaatga	actgcaggac	gaggcagcgc	3900
ggctatcgtg	gctggccacg	acgggcgttc	cttgccgcagc	tgtgctcgac	gttgtcactg	3960
aagcgggaag	ggactggctg	ctattgggcg	aagtgccggg	gcaggatctc	ctgtcatctc	4020
accttgctcc	tgccgagaaa	gtatccatca	tggctgatgc	aatgcggcgg	ctgcatacgc	4080
ttgatccggc	tacctgcccc	ttcgaccacc	aagcgaacaa	tgcacatcgag	cgagcacgta	4140
ctcggatgga	agccgggtct	gtcgatcagg	atgatctgga	cgaagagcat	caggggctcg	4200
cgccagccga	actgttcgcc	aggctcaagg	cgccatgcc	cgacggcgag	gatctcgtcg	4260
tgacccatgg	cgatgcctgc	ttgccgaata	tcatgggtga	aaatggccgc	ttttctggat	4320
tcacgactg	tggccggctg	ggtgtggcgg	accgctatca	ggacatagcg	ttggctaccc	4380
gtgatattgc	tgaagagctt	ggcggcgaa	gggctgaccg	cttcctcgtg	ctttacggta	4440
tgcgcgctcc	cgattcgcag	cgcacgcct	tctatcgctt	tcttgacgag	ttcttctgag	4500
cgggactctg	gggttcgaaa	tgaccgacca	agcgacgccc	aacctgccat	cacgagattt	4560
cgattccacc	gccgccttct	atgaaagggt	gggcttcgga	atcgtttttc	gggacgcccg	4620
ctggatgatc	ctccagcgcg	gggatctcat	gctggagttc	ttcgccacc	ccaacttggt	4680
tattgcagct	tataatgggt	acaaataaag	caatagcatc	acaaatttca	caaataaagc	4740
atthtttttca	ctgcattcta	gttgtgggtt	gtccaaactc	atcaatgtat	cttatcatgt	4800
ctgtataccg	tcgacctcta	gctagagctt	ggcgtaatca	tggctatagc	tgtttcctgt	4860
gtgaaattgt	tatccgctca	caattccaca	caacatacga	gccggaagca	taaagtgtaa	4920

```

agcctgggggt gcctaatagag tgagctaact cacattaatt gcgttgcgct cactgcccgc 4980
tttccagtcg ggaaacctgt cgtgccagct gcattaatga atcggccaac gcgcggggag 5040
aggcggtttg cgtattgggc gctcttcgc ttcctcgtc actgactcg tgcgctcgg 5100
cgttcggtg cggcgagcgg tatcagctca ctcaaaggcg gtaatacgg tatccacaga 5160
atcaggggat aacgcaggaa agaacatgtg agcaaaagg cagcaaaagg ccaggaaccg 5220
taaaaaggcc gcgttgctgg cgtttttcca taggctccgc cccctgacg agcatcacia 5280
aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga ctataaagat accaggcggt 5340
tccccctgga agctccctcg tgcgctctcc tgttccgacc ctgccgctta ccggatacct 5400
gtccgccttt ctcccttcgg gaagcgtggc gctttctcat agctcacgct gtaggtatct 5460
cagttcggtg taggtcggtt gctccaagct gggctgtgtg cacgaacccc ccgttcagcc 5520
cgaccgctgc gccttatccg gtaactatcg tcttgagtc aacccggtta gacacgactt 5580
atcgccactg gcagcagcca ctggtaacag gattagcaga gcgaggtatg taggcggtgc 5640
tacagagttc ttgaagtggg ggcctaacta cggctacact agaagaacag tatttggtat 5700
ctgcgctctg ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt gatccggcaa 5760
acaaaccacc gctggtagcg gtttttttgt ttgcaagcag cagattacgc gcagaaaaaa 5820
a

```

<210> 41

<211> 10075

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic recombinant DNA molecule sequence

<220>

<221> modified\_base

<222> (9040)..(9040)

<223> unsequenced nucleotide

<400> 41

```

ctggtagcgg tggttttttt gtttgcaagc agcagattac gcgcagaaaa aaaggatctc 60
aagaagatcc tttgatcttt tctacggggg ctgacgctca gtggaacgaa aactcacgtt 120
aagggtttt ggtcatgaga ttatcaaaaa ggatcttcac ctagatcctt ttaaattaaa 180
aatgaagttt taaatcaatc taaagtatat atgagtaaac ttggtctgac agttaccaat 240
gcttaatcag tgaggcacct atctcagcga tctgtctatt tcgttcaccc atagttgcct 300
gactccccgt cgtgtagata actacgatac gggagggtt accatctggc cccagtgtg 360
caatgatacc gcgagaccca cgctcacggg ctccagattt atcagcaata aaccagccag 420
cgggaagggg cgagcgcaga agtgggtcct caactttatc cgctccatc cagtctatta 480
attgttgccg ggaagctaga gtaagtagtt cgccagttaa tagtttgcg aacgttggtg 540
ccattgctac aggcacgtg gtgtcacgct cgtcgtttgg tatggcttca ttcagctccg 600
gttcccaacg atcaaggcga gttacatgat ccccatgtt gtgcaaaaaa gcggttagct 660
ccttcggtcc tccgatcggt gtcagaagta agttggccgc agtgttatca ctcatggtta 720
tggcagcact gcataattct ctactgtca tgccatccgt aagatgcttt tctgtgactg 780
gtgagtactc aaccaagtca ttctgagaat agtgtatcgc gcgaccgagt tgctcttgcc 840
cggcgtcaat acgggataat accgcgccac atagcagaac tttaaaagtg ctcatcattg 900
gaaaacgttc ttcggggcga aaactctcaa ggatcttacc gctgttgaga tccagttcga 960
tgtaaccac tcgtgcaccc aactgatctt cagcatcttt tactttcacc agcgtttctg 1020
ggtgagcaaa aacaggaagg caaaatgccg caaaaaaggg aataaggcg acacggaaat 1080
gttgaatact catactcttc ctttttcaat attattgaag catttatcag ggttattgtc 1140
tcatgagcgg atacatattt gaatgtattt agaaaaataa acaaataagg gttccgcgca 1200
catttccccg aaaagtgcc cctaaattgt aagcgttaat attttgtaa aattcgcgtt 1260
aaatttttgt taaatcagct cattttttaa ccaataggcc gaaatcggca aaatccctta 1320
taaatcaaaa gaatagaccg agatagggtt gagtggtgtt ccagtttgga acaagagtcc 1380
actattaaag aacgtggact ccaacgtcaa agggcgaaaa accgtctatc agggcgatgg 1440
cccactacgt gaaccatcac cctaatacag ttttttgggg tcgaggtgcc gtaaagcact 1500

```



aaatcggaac	cctaaagggg	gcccccgatt	tagagcttga	cggggaaagc	cggcgaacgt	1560
ggcgagaaa	gaaggggaag	aagcgaaaag	agcgggcgct	agggcgctgg	caagtgtagc	1620
ggtcacgctg	cgcgtaacca	ccacacccgc	cgcgcttaat	gcgcgcgtac	agggcgcgctc	1680
ccattcgcca	ttcaggctgc	gcaactgttg	ggaagggcga	tcgggtgcggg	cctcttcgct	1740
attacgccag	ctggcgaaa	ggggatgtgc	tgcaaggcga	ttaaagtggg	taacgccagg	1800
gttttccag	tcacgacgtt	gtaaaacgac	ggccagttag	cgcgcgtaat	acgactcact	1860
atagggcgaa	ttggagctcc	accgcggtgc	ggccggggcca	tgaggccac	gacatgataa	1920
gatacattga	tgagtttgga	caaaccacaa	ctagaatgca	gtgaaaaaaa	tgctttat	1980
gtgaaatttg	tgatgctatt	gctttat	taaccattat	aagctgcaat	aaacaagttc	2040
ttgtacagct	cgtccatgcc	gagagtgtac	cgggcggcgg	tcacgaactc	cagcaggacc	2100
atgtgatcgc	gcttctcggt	ggggtc	ctcagggcgg	actggtagct	caggtagtgg	2160
ttgtcgggca	gcagcacggg	ggcgtcgcg	atgggggtgt	tctgctggta	gtggtcggcg	2220
agctgcacgc	tgccgtcctc	gatgttgtgg	cggatcctga	agttcacctt	gatgccgttc	2280
ttctgcttgt	cggccatgat	atagacgttg	tggtgttgt	agttgtactc	cagcttgtgc	2340
cccaggatgt	tgccgtcctc	cttgaagtcg	atgcccttca	gctcgatgcg	gttcaccagg	2400
gtgtcgccct	cgaacttcac	ctcggcgcg	gtctgttagt	tgccgtcgtc	cttgaagaag	2460
atggtcgct	cctggacgta	gcttcgggc	atggcgact	tgaagaagtc	gtgctgcttc	2520
atgtggtcgg	ggtagcgggc	gaagcactgc	aggcgcgtagc	cgaaggtggg	cacgaggggtg	2580
ggccagggca	cgggcagctt	ggcggtgtgt	cagatgaact	tcagggtcag	cttgccgtag	2640
gtggcatcgc	cctcgccctc	ggcgacacg	ctgaacttgt	ggcgtttac	gtcgccgtcc	2700
agctcgacca	ggatgggcac	caccccggtg	aacagctcct	cgcccttgct	caccatgggtg	2760
gcgttagctt	gatttgacag	tggctggggg	ttgcgcgcgc	gggttttata	ggaagccaca	2820
gcggccactc	gagccataaa	aggcaacttt	aggaacggcg	gggggtgatt	ggattcgagt	2880
cgtttatcca	ccggccttgc	cgcacagtgc	agcatttttt	tacccctct	cccctccttt	2940
tgcgggggaa	aaaaaaaaaa	aaaaaaaaaa	aaggagaaga	gaaaaaaagc	gagcgagaga	3000
gaaagcgaga	ttgaggaaga	ggatgaagag	tttggcgatg	ggtgctgggt	ccgtaggccc	3060
agatggacaa	gaatagcccc	cgcccttgcg	gacagtatcc	cattcagtga	ctcagatcag	3120
atcaagcggc	cgccagtggt	atggatatct	gcagaattcc	agcacactgg	cggccgttac	3180
tagtggatcc	ctctggagct	gcctagtgtg	ggtagctgct	ggaaacttct	aaaatgtctc	3240
ggcctattta	attggagaat	tcactgaaga	attaaaggac	tcagtaaaga	agaggtttgg	3300
tttgttcag	aaagtaagta	gcttgtgaaa	tgctttat	tcctgtgact	aaaattccac	3360
cttgtaattc	ccaataagga	ttgcttcaag	ctattttaata	caaattaaac	ctgtttcacc	3420
atgctcatag	gaaaatgtca	aagtatcttc	attcaccaac	tgagggctgg	ctcttctttg	3480
aaggtagctc	tgagatttgt	tttgtaattg	agacaaagac	ttatttaact	gttggggagg	3540
gggaaggagg	acttagggcca	ggcaggcggg	gatgtggggg	acgaaccatg	aggggcaatt	3600
tggtcgctcc	tgaaaagggtc	cagagctacc	ttaaaagaca	ggctttctat	aacaagtggg	3660
cctttctttt	cctttctttt	tttttttgag	acagagtctc	actctgttgc	ccaagctgga	3720
gtgcagtggt	gcgatctcag	ctcactgcaa	cctccagctc	ccaggttcaa	gcaattctcc	3780
tgctcagcc	tcccaagtag	ctgggattac	aggcaaccgc	taccatgccc	agctaatttt	3840
ttgtat	aatagagatg	gggtttcacc	atgttggcca	gggtggtctc	aaactcctga	3900
cctcaagtga	tccacccacc	tcagcctccc	aaagtgtctg	gattacagac	atgagccacc	3960
gcgcccagac	acagctgggt	ttgtcgccca	ggctggaagt	gcagtggcac	gatctcggt	4020
cactgcaagc	tctgcctcct	gggttcacgc	cattctcctg	cctcagcttc	ccaagtagct	4080
gagactacag	gcgcccggcca	ccacgcccgg	ctaatttttt	gtatttttag	tgagatgggt	4140
gtttcaccat	gttagccagg	atggtcttga	tctcctgacc	tcgtgatccg	cctgtcttgg	4200
cctcccaaag	tgctgggatg	acaggcggtg	gccaccacgc	ctggcttaca	agtggccctt	4260
tctaagcccc	aagatgtcac	tgtcttttca	ctgaaggcgg	aaaaagaatg	tagggcatgtt	4320
tgtctgagag	ggaggggagta	gttttaggag	aaaggaccaa	gagcacattt	tgagtgggaa	4380
gtaagtgggt	tccttttcta	gccttatgat	atggtttggc	tctgtgtccc	cacccaaatc	4440
tcatggtgaa	ctgtaatccc	cagtgggtgga	ggtgaggcct	ggcgggaggg	actggatcat	4500
gggggtgatc	cttcatgaat	ggtttagcac	catcctcttg	gtgctactct	agtgatagcg	4560
ttctcacaa	atctgatggg	tgaaaagtgt	gtgtctcttg	tttttgcctc	tgccatgcaa	4620
gacgcctact	cctccttcac	cttctgtcat	gagtcacaatt	ccctgaggcc	tcccagaag	4680
cagaagcttc	ctgtacagcc	ttgtggggcca	cgagccaatt	aaacctcttt	tctttataaa	4740
ttaccaggtc	tcagggattt	ctttatagca	gactaataca	cctcaaagta	gggatttatt	4800
ttattattct	at	tgagataggg	tctcgatctg	tcaccagggc	tgagtgtagc	4860
tggtcatgatc	tcagcttgct	gtaacctctg	cctcccggtt	tgaagcaatt	ctcatgccta	4920
agccacccaa	gtagctggaa	ttacaggtgt	gtgccaccct	gctcagctaa	tttttgcata	4980

tttagtagag	ttgggggtttc	tccatgttgg	ccaggctggt	atcaaactcc	tggactcaag	5040
caatctgccc	accttggcct	ccccagagt	ccaggattac	aggcatgacc	catcttgcct	5100
ggtcagtttt	ttttttttta	aatccagtta	atgctgggtg	ccatgactca	cacctgtaat	5160
gccagtgcct	tgggaggccg	aggtgggaag	atcacttgag	gccaggagtt	tgagaccagc	5220
ctgggaaaca	tctcatctct	acaaaaaac	aaaaacaaaa	caaaacaaca	acaacatcaa	5280
acatagccag	gtgtggtggt	gcacacctgt	agtcccagct	ccttgggagg	ctgaggcagg	5340
aggattgctt	gagcccagga	gttcaaggct	gcagtgaact	atgattgcac	cactgcattc	5400
cagcctgggc	aacagagcaa	gaacctatct	ctattttaaga	aaaaatgctc	ctaagaaatc	5460
atcttagtac	ttggtaagct	attaagggtt	tttttaagt	tattaattac	atagccagta	5520
taataactta	gttgctaaga	gctttgaggt	ggctccaagc	aagaagaaa	gagttgttgt	5580
taagctacag	ctcagagtct	cagaatgaac	tgtttccatc	ttcgctgtcc	tttgcttaac	5640
tgtgggataa	atataatttt	ctcaacattc	aacacatgtg	tccttaatcg	tcctctacga	5700
cactggttcc	atgttgccaa	gagaaccag	aacagccgca	aacacctact	ggattgttac	5760
tggcctttcc	tgagaggatg	acccgggctt	tgacctgtgt	catgttgaag	ttcttcaagt	5820
aggcctcgta	gattctcttg	gccagagatt	tgagatctgc	agtttcagaa	tcttctatgt	5880
catgttcaca	ggttaagaatt	tctgctttca	gttttgcttt	ctcagatcct	ggcattcgct	5940
cggtagccgg	cgatacagtc	aactgtcttt	gacctttgtt	actactctct	tccgatgatg	6000
atgtcgcact	tattctatgc	tgtctcaatg	ttagaggcat	atcagtctcc	actgaagcca	6060
atctatctgt	gacggcatct	ttattcacat	tatcttgtac	aaataatcct	gttaacaatg	6120
cttttatatc	ctgtaaagaa	tccattttca	aaatcatgtc	aaggctcttct	cgaggaaaaa	6180
tcagtagaaa	tagctgttcc	agtcttttca	gccttgattc	cacttctgtc	agatgtgccc	6240
tagtcagcgg	agaccttttg	gttttgggag	agttagcgaca	ctcccagttg	ttcttcagac	6300
acttggcgca	cttcgggtttt	cttttggagc	acttgagctt	tttaagtctg	caaatatcgc	6360
atgcttgctt	gatagaagac	agtagcttca	tggtaggcaag	cttaagttta	aacgctagaa	6420
ctagtggatc	cggataagcc	agtaagcagt	gggttctcta	gttagccaga	gagctctgct	6480
tatatagacc	tcccaccgta	cacgcctacc	gcccatttgc	gtcaatgggg	cggagtgtgt	6540
acgacatttt	ggaaagtccc	gttgattttg	gtgccaaaac	aaactcccat	tgacgtcaat	6600
gggtgggaga	cttggaatac	ccgtgagtc	aaaccgctat	ccacgcccac	tgatgtactg	6660
ccaaaaccgc	atcaccatgg	taatagcgat	gactaatacg	tagatgtact	gccaaagtag	6720
aaagtcccat	aaggctcatgt	actgggcata	atgccaggcg	ggccatttac	cgctattgac	6780
gtcaataggg	ggcgactctg	gcatatgata	cacttgatgt	actgccaagt	gggcagttta	6840
ccgtaaatac	tccacccatt	gacgtcaatg	gaaagtccct	attggcgcta	ctatgggaac	6900
atacgtcatt	attgacgtca	atggggcggt	gtcgttgggc	ggtcagccag	gcgggccatt	6960
taccgtaagt	tatgtaacgc	ggaactccat	atatgggcta	tgaactaatg	accccgtaat	7020
tgattactat	taataactag	tcaataatca	atgtcaaccc	cgggctgcag	gaattctacc	7080
gggtagggga	ggcgcttttc	ccaaggcagt	ctggagcatg	cgcttttagca	gccccgctgg	7140
cacttggcgc	atcacaagt	gcctctggcc	tcgcacacat	tccacatcca	ccggtagcgc	7200
caaccggctc	ccttcttttg	tggccccctc	gcgccacctt	ctactcctcc	cctagtcaag	7260
aagtcccccc	ccgccccgca	gctcgcgctg	tgaggagcgt	gacaaatgga	agtagcacgt	7320
ctcactagtc	tcgtgcagat	ggacaagcag	cgctgagcaa	tggaaagcgg	taggcctttg	7380
gggcagcgcc	caatagcagc	ttggctcctt	cgctttctgg	gctcagaggc	tgggaagggg	7440
tgggtccggg	ggcgggctca	ggggcgggct	cagggcgggg	gcgggcgcga	aggtcctccg	7500
gaccggcat	tctgcacgct	tcaaaagcgc	acgtctgccg	cgctgttctc	ctcttcctca	7560
tctccgggcc	ttcgacctgc	atgaaaaagc	ctgaactcac	cgcgacgtct	gtcgagaagt	7620
ttctgatcga	aaagttcgac	agcgtctccg	acctgatgca	gctctcgag	ggcgaagaat	7680
ctcgtgcttt	cagcttcgat	gtaggagggc	gtggatatgt	cctgcgggta	aatagctgcg	7740
ccgatgggtt	ctacaaagat	cgttatgttt	atcggcactt	tgcatcggcc	gcgctcccga	7800
ttccggaagt	gcttgacatt	ggggaattca	gcgagagcct	gacctattgc	atctcccgcc	7860
gtgcacaggg	tgtcacgttg	caagacctgc	ctgaaaccga	actgcccgtc	gttctgcagc	7920
cggtcgcgga	ggccatggat	gcgatcgctg	cggccgatct	tagccagacg	agcgggttcg	7980
gcccattcgg	accgcaagga	atcgggcaat	acactacatg	gcgtgatttc	atatgcgcga	8040
ttgtgatccc	ccatgtgtat	cactggcaaa	ctgtgatgga	cgacaccgtc	agtgcgtccg	8100
tcgcgcagcg	tctcgatgag	ctgatgcttt	gggcccagga	ctgccccgaa	gtccggcacc	8160
tcgtgcacgc	ggatttcggc	tccaacaatg	tcctgacgga	caatggccgc	ataacagcgg	8220
tcattgactg	gagcgaggcg	atgttcgggg	attcccaata	cgaggtcgcc	aacatcttct	8280
tctggaggcc	gtggttggct	tgtatggagc	agcagacgcg	ctacttcgag	cggaggcatc	8340
cggagcttgc	aggatcgccg	cggctccggg	cgtatatgct	ccgcattggg	cttgaccaac	8400
tctatcagag	cttggttgac	ggcaatttct	atgatgcagc	ttgggcgcag	ggtcgatgcg	8460

acgcaatcgt	ccgatccgga	gccgggactg	tccggcgctac	acaaatcgcc	cgcagaagcg	8520
cggccgtctg	gaccgatggc	tgtgtagaag	tactcgccga	tagtggaaac	cgacgccccca	8580
gcactcgtcc	gagggcaaa	gaatgcctga	gaaaggaagt	gagctgtaaa	ggctgagctc	8640
tctctctgac	gtatgtagcc	tctggttagc	ttcgtcactc	actgttcttg	actcagcatg	8700
gcaatctgat	gaaatcccag	ctgtaagtct	gcataaattg	atgatctatt	aaacaataaa	8760
gatgtccact	aaaatggaag	ttttttactg	tcatactttg	ttaagaaggg	tgagaacaga	8820
gtacctacat	tttgaatgga	aggattggag	ctacgggggt	gggggtgggg	gtgggattag	8880
ataaatgcct	gctctttact	gaaggctctt	tactattgct	ttatgataat	gtttcatagt	8940
tggatatcat	aattttaaaca	agcaaaaacca	aattaagggc	cagctcattc	ctccactcat	9000
gatctataga	tctatagatc	tctcgtggga	tcattgtttt	tctgatccac	tggaagctta	9060
tcgataccgt	cgacctcgag	gggggggccc	gtaccagct	tttgttccct	ttagtggagg	9120
ttaattgcgc	gcttggcgta	atcatgggtca	tagctgtttc	ctgtgtgaaa	ttgttatccg	9180
ctcacaattc	cacacaacat	acgagccgga	agcataaagt	gtaaagcctg	gggtgcctaa	9240
tgagtgaagt	aactcacatt	aattgcgttg	cgctcactgc	ccgctttcca	gtcgggaaac	9300
ctgtcgtgcc	agctgcatta	atgaatcggc	caacgcgcgg	ggagaggcgg	tttgcgattt	9360
gggcgctcct	ccgcttcttc	gctcactgac	tcgctgcgct	cggtcgttcg	gctgcggcga	9420
gcggtatcga	ctcactcaaa	ggcggttaata	cgggtatcca	cagaatcagg	ggataacgca	9480
ggaaagaaca	tgtgagcaaa	aggccagcaa	aaggccagga	accgtaaaaa	ggccgcgttg	9540
ctggcgtttt	tccataggct	ccgccccct	gacgagcatc	acaaaaatcg	acgctcaagt	9600
cagaggtggc	gaaacccgac	aggactataa	agataccagg	cgtttcccc	tggaagctcc	9660
ctcgtgcgct	ctcctgttcc	gacctgccc	cttaccggat	acctgtccgc	ctttctccct	9720
tcgggaagcg	tggcgctttc	tcatagtcca	cgctgtaggt	atctcagttc	gggtgaggtc	9780
gttcgctcca	agctgggctg	tgtgcacgaa	ccctcggttc	agcccgaccg	ctgcgcctta	9840
tccggttaact	atcgtcttga	gtccaaccgc	gtaagacacg	acttatcgcc	actggcagca	9900
gccactggta	acaggattag	cagagcgagg	tatgtaggcg	gtgctacaga	gttcttgaag	9960
tggtggccta	actacggcta	cactagaagg	acagtatttg	gtatctgcgc	tctgctgaag	10020
ccagttacct	tcggaaaaag	agttggtagc	tcttgatccg	gcaaaacaaac	caccg	10075

&lt;210&gt; 42

&lt;211&gt; 8971

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Synthetic recombinant DNA vector sequence

&lt;400&gt; 42

gtaggtgtca	ttctattctg	gggggtgggg	tggggcagga	cagcaagggg	gaggattggg	60
aagacaatag	caggcatgct	ggggatgcgg	tgggctctat	ggcttctgag	gcggaaagaa	120
ccagctgggg	ctctaggggg	tatccccacg	cgccctgtag	cggcgcatta	agcgcggcgg	180
gtgtggtggt	tacgcgcagc	gtgaccgcta	cacttgccag	cgccctagcg	cccgtcctt	240
tcgctttctt	cccttccttt	ctcgccacgt	tcgcggcgtt	tccccgtcaa	gctctaaatc	300
gggggtctcc	tttagggttc	cgatttagtg	ctttacggca	cctcgacccc	aaaaaacttg	360
attagggtga	tggttcacgt	agtgggcat	cgccctgata	gacggttttt	cgccctttga	420
cgttggagtc	cacgttcttt	aatagtgagc	tcttggtcca	aactggaaca	acactcaacc	480
ctatctcggt	ctattctttt	gatttataag	ggattttgcc	gatttcggcc	tattgggttaa	540
aaaatgagct	gatttaacaa	aaatttaacg	cgaattaatt	ctgtggaatg	tgtgtcagtt	600
agggtgtgga	aagtccccag	gctccccagc	aggcagaagt	atgcaaagca	tgcattctcaa	660
ttagtcagca	accaggtgtg	gaaagtcccc	aggctcccca	gcaggcagaa	gtatgcaaag	720
catgcatctc	aattagtcag	caaccatagt	cccgcctcta	actccgcccc	tcccgccctt	780
aactccgccc	agttccgccc	attctccgcc	ccatggctga	ctaatttttt	ttattttatgc	840
agagggccgag	gccgcctctg	cctctgagct	attccagaag	tagtgaggag	gctttttttg	900
aggcctaggc	ttttgcaaaa	agctcccggg	agcttgata	tccatttttc	gatctgatca	960
gcacgtgatg	aaaaagcctg	aactcaccgc	gacgtctgtc	gagaagtttc	tgatcgaaaa	1020
gttcgacagc	gtctccgacc	tgatgcagct	ctcggagggc	gaagaatctc	gtgctttcag	1080
cttcgatgta	ggagggcggtg	gatatgtcct	gcgggtaaat	agctgcgccg	atggtttcta	1140

caaagatcgt	tatgtttatc	ggcactttgc	atcggccgcg	ctcccgaattc	cggaagtgc	1200
tgacattggg	gaattcagcg	agagcctgac	ctattgcac	tcccgcctg	cacagggtgt	1260
cacgttgcaa	gacctgcctg	aaaccgaact	gcccgcgtgt	ctgcagccgg	tcgcggaggc	1320
catggatgcg	atcgctgcgg	ccgatcttag	ccagacgagc	gggttcggcc	cattcggacc	1380
gcaaggaatc	ggccaataca	ctacatggcg	tgatttcata	tgcgcgattg	ctgatcccca	1440
tgtgtatcac	tggcaaaactg	tgatggacga	caccgtcagt	gcgtccgtcg	cgcaggctct	1500
cgatgagctg	atgctttggg	ccgaggactg	ccccgaagtc	cggcacctcg	tgacgcgga	1560
tttcggctcc	aacaatgtcc	tgacggacaa	tggccgcata	acagcgggtca	ttgactggag	1620
cgaggcgatg	ttcggggatt	cccaatacga	ggtcgccaac	atcttcttct	ggaggccgtg	1680
gttggcttgt	atgggacgc	agacgcgcta	cttcgagcgg	aggcatccgg	agcttgacgg	1740
atcggccggg	ctccggggcg	atatgctccg	cattgggtctt	gaccaactct	atcagagctt	1800
ggttgacggc	aatttctgatg	atgcagcttg	ggcgcagggt	cgatgcgacg	caatcgctcg	1860
atccggagcc	gggactgtcg	ggcgtaacac	aatcgcccg	agaagcgcg	ccgtctggac	1920
cgatggctgt	gtagaagtac	tcgcccgatag	tggaaaccga	cgcgccagca	ctcgcccgag	1980
ggcaaaggaa	tagcacgtgc	tacgagatctt	cgattccacc	gccgccttct	atgaaagggt	2040
gggcttcgga	atcggttttcc	gggacgcggg	ctggatgac	ctccagcgcg	gggatctcat	2100
gctggagttc	ttcgcccacc	ccaacttggt	tattgcagct	tataatgggt	acaaataaag	2160
caatagcatc	acaaatttca	caaataaagc	atcttttttca	ctgcattcta	gttgtgggtt	2220
gtccaaactc	atcaatgtat	cttatcatgt	ctgtataccg	tcgacctcta	gctagagctt	2280
ggcgtaatca	tgggtcatagc	tgtttcctgt	gtgaaattgt	tatccgctca	caattccaca	2340
caacatacga	gccggaagca	taaagtgtaa	agcctggggg	gcctaattgag	tgagctaact	2400
cacattaact	gcgttgccgt	cactgcccgc	tttccagtcg	ggaaacctgt	cgtagccagt	2460
gcattaatga	atcggccaac	gcgcggggag	aggcgggttg	cgtattgggc	gctcttccgc	2520
ttcctcgctc	actgactcgc	tgcgctcggt	cgttcggtcg	cggcgagcgg	tatcagctca	2580
ctcaaaggcg	gtaatacggg	tatccacaga	atcaggggat	aacgcaggaa	agaacatgtg	2640
agcaaaaggc	cagcaaaagg	ccaggaaccg	taaaaaggcc	gcgttgctgg	cgtttttcca	2700
taggctccgc	ccccctgacg	agcatcacaa	aaatcgacgc	tcaagtcaga	ggtggcgaaa	2760
cccgacagga	ctataaagat	accaggcggt	tccccctgga	agctccctcg	tgcgctctcc	2820
tgttccgacc	ctgcgcgtta	ccggatacct	gtccgccttt	ctcccttcgg	gaagcgtggc	2880
gctttctcat	agctcacgct	gtaggtatct	cagttcggtg	taggtcggtc	gctccaagct	2940
gggctgtgtg	cacgaacccc	ccgttcagcc	cgaccgctgc	gccttatccg	gtaactatcg	3000
tcttgagtcc	aacccggtaa	gacacgactt	atcgccactg	gcagcagcca	ctggtaacag	3060
gattagcaga	gcgagggtatg	taggcgggtgc	tacagagttc	ttgaagtggg	ggcctaacta	3120
cggctacact	agaagaacag	tatttggtat	ctgcgctctg	ctgaagccag	ttacctctcg	3180
aaaaagagtt	ggtagctctt	gatccggcaa	acaaaccacc	gctggtagcg	gtttttttgt	3240
ttgcaagcag	cagattacgc	gcagaaaaaa	aggatctcaa	gaagatcctt	tgatcttttc	3300
tacggggctc	gacgctcagt	ggaacgaaaa	ctcacgttaa	gggattttgg	tcatgagatt	3360
atcaaaaagg	atcttcacct	agatcctttt	aaattaaaaa	tgaagtttta	aatcaatcta	3420
aagtatatat	gagtaaaactt	ggctctgacag	ttaccaatgc	ttaatcagtg	aggcacctat	3480
ctcagcgatc	tgtctatctt	gttcatccat	agttgcctga	ctcccgcctg	tgtagataac	3540
tacgatacgg	gagggcttac	catctggccc	cagtgtcgca	atgataccgc	gagacccacg	3600
ctcaccggct	ccagatttat	cagcaataaa	ccagccagcc	ggaaggggccg	agcgcagaag	3660
tggctctgca	actttatccg	cctccatcca	gtctattaat	tggtgccggg	aagctagagt	3720
aagtagttcg	ccagtttaata	gtttgcgcaa	cgttgttgcc	attgctacag	gcacgtgggt	3780
gtcacgctcg	tcgttttggt	tggcttcatt	cagctccggt	tcccaacgat	caaggcgagt	3840
tacatgatcc	cccatggtgt	gcaaaaaagc	ggttagctcc	ttcggtcctc	cgatcgttgt	3900
cagaagtaag	ttggccgcag	tgttatcact	catggttatg	gcagcactgc	ataattctct	3960
tactgtcatg	ccatccgtaa	gatgcttttc	tgtgactggg	gagtactcaa	ccaagtcatt	4020
ctgagaatag	tgtatgcggc	gaccgagttg	ctcttgcccg	gcgtcaatac	gggataatac	4080
cgcgccacat	agcagaactt	taaaagtgtc	catcattgga	aaacgttctt	cggggcgaaa	4140
actctcaagg	atcttaccgc	tggttgagatc	cagttcgatg	taaccactc	gtgcacccaa	4200
ctgatcttca	gcatctttta	ctttcaccag	cgtttctggg	tgagcaaaaa	caggaaggca	4260
aaatggcgca	taagggggaa	taaggggcag	acggaaatgt	tgaatactca	tactcttctc	4320
ttttcaatat	tattgaagca	tttatcaggg	ttattgtctc	atgagcggat	acatatattga	4380
atgtatttag	aaaaataaac	aaataggggt	tccgcgcaca	tttccccgaa	aagtgccacc	4440
tgacgtcgac	ggatcgggag	atctcccgat	cccctatggg	gcactctcag	tacaatctgc	4500
tctgatgccg	catagttaag	ccagtatctg	ctccctgctt	gtgtgttgga	ggtcgctgag	4560
tagtgcgcca	gcaaaaattta	agctacaaca	aggcaaggct	tgaccgacaa	ttgcatgaag	4620

aatctgctta	gggttaggcg	ttttgcgctg	cttcgcgatg	tacggggccag	atatacgcgt	4680
tgacattgat	tattgactag	ttattaatag	taatcaatta	cgggggtcatt	agttcatagc	4740
ccatatatgg	agttccgcgt	tacataactt	acggtaaatg	gcccgcctgg	ctgaccgccc	4800
aacgaccccc	gccatttgac	gtcaataatg	acgtatgttc	ccatagtaac	gccaataggg	4860
actttccatt	gacgtcaatg	ggtggagtag	ttacggtaaa	ctgcccactt	ggcagtacat	4920
caagtgtatc	atatgccaa	tacgccccct	attgacgtca	atgacggtaa	atggcccgcc	4980
tggcattatg	cccagtacat	gaccttatgg	gactttccta	cttggcagta	catctacgta	5040
ttagtcatcg	ctattaccat	ggtgatggcg	ttttggcagt	acatcaatgg	gcgtggatag	5100
cggtttgact	cacggggatt	tccaagtctc	caccccattg	acgtcaatgg	gagtttgttt	5160
tggcaccaaa	atcaacggga	ctttccaaaa	tgtcgtaaca	actccgcccc	attgacgcaa	5220
atgggcggta	ggcgtgtacg	gtgggagggtc	tatataagca	gagctctctg	gctaactaga	5280
gaaccctactg	cttactgggt	tatcgaaatt	aatacgactc	actataggga	gacccaagct	5340
ggctagcgtt	taaacttaag	cttgccacca	tgaagctact	gtcttctatc	gaacaagcat	5400
gcgatatttg	ccgacttaaa	aagctcaagt	gtcctcaaga	aaaaccgaag	tgcgccaagt	5460
gtctgaagaa	caactgggag	tgtcgctact	ctcccaaaac	caaaaggctc	ccgctgacta	5520
gggcacatct	gacagaagtg	gaatcaaggc	tagaaagact	ggaacagcta	tttctactga	5580
ttttcctcgc	agaagacctt	gacatgattt	tgaaaatgga	ttctttacag	gatataaaag	5640
cattgttaac	aggattattt	gtacaagata	atgtgaataa	agatgccgtc	acagatagat	5700
tggcttcagt	ggagactgat	atgcctctaa	cattgagaca	gcatagaata	agtgcgacat	5760
catcatcgga	agagagtagt	aacaaagggtc	aaagacagtt	gactgtatcg	ccgggtaccg	5820
cagtggggat	gtctcataat	ggtaagtaaa	cagtcatcac	catatacttt	attattctca	5880
ttatagctgc	cagaccagtg	gacactaaag	ccattgccaa	aaatgtgtac	agttttttcca	5940
ccaaatgccca	gaatttagaa	tattgcatgg	cgataaaaca	tttctctttt	aggtcagtgt	6000
ttttaaagtt	ttattataga	acctttctct	ctgtgggttg	gcatctgcca	tgaggagaaa	6060
agagacttga	aaaatctggg	ggattatggg	aaaaacttta	gtgttagaac	aaatataatg	6120
atcatcatga	ttaattttta	tattgcctca	tgcattaaaa	acagaattaa	aattttacttt	6180
tgtgtatcag	tagtgtcatt	ctgcttttaa	tgtaaaaaaa	tcctgaagat	gtctgggtct	6240
gaaactttaa	atggcatatg	agaatttaga	ttcatttaat	tttttttttt	tttttttttt	6300
tgtgttttat	tatgcaaaaga	cgagtattcg	tttccttcac	ttcaagggaac	gtccatgagt	6360
tgtggaaagg	gtaaaaacaaa	ttatttttat	agttctgcaa	aacttcaaag	ccatagaaac	6420
ataagagaat	aaaactttct	acttttgaaa	cctgtgtcta	aagtcaagct	tatgactttg	6480
ggtgtcaact	taaggggaaat	gatgtagaag	acctgagaaa	agacccatgt	tttcatgatc	6540
tgtggaacta	actcagtaag	gtcctggaag	acaaacgtgt	ttgaaaaaat	aattcttggt	6600
cactatttgg	ttctgattgg	ttgcctggct	gacaaaaata	attttagggt	taacttgata	6660
gtattatag	tgtctagtag	taatcagtta	ttttcaatgc	actcaaggaa	taaaaaatgc	6720
caacatgatc	aaattaaatt	tgaagtttca	tttcatcagc	tatcctattt	tttctttctt	6780
ttgttttcat	ctctctcaaa	actgatctag	taagaaggca	gtcttcatct	atatgtactc	6840
taaagcatct	gagaaaaatta	aatttttgct	ctatcatgca	gatacttggt	tttcacagtg	6900
gaaaccatca	agatatattga	gagtcactca	ttgattaata	acctaccctt	taaaaaatagt	6960
tcacaaaaat	ctatctgtat	gattactaat	gggtcagttc	ctctaaaggc	aagaatgtat	7020
tttgaataga	gctatcagca	gtactgaaca	ggatggaaat	ttggtccaac	tttcttcaac	7080
aattgactac	atctgctgta	tacaactgta	gtcatggatt	gcaaaataat	ttgctggttc	7140
agttccatca	gtttcctttt	tatttcagac	ctcataaaac	ataagggcaa	gcttccaagt	7200
caagggttgt	attttaggtg	atgcaatcaa	gttgcaataa	acttatttca	tcttcatgaa	7260
gaaatcattg	actttgaggt	cattatcact	gagacagcct	aggaaagcat	gccagttccg	7320
cttttaacac	taatttttgc	ctattaaaaa	attatcactt	cacacattca	tgtaaaaatat	7380
aattcctggt	acttacaggc	agtgcataca	ttgatgtaaa	atatatttcc	tgttgcttac	7440
aggcattgca	tacgtttact	taaaatcaat	ctgcaattgc	ttattgttta	taggcaaata	7500
ttatttaaac	cccaacctat	gtagtatgga	tgggtataatt	ttctcctccc	ataaatgtca	7560
aagaaagcta	tcataaactt	gatagggtcat	attttgactt	catatgacaa	ttatttttat	7620
ttctagagag	ggtcaatctt	gataggggagt	gtttgttcag	ttctagcaaa	agcaagaaaa	7680
tatatgaatt	aagtgtcatg	acatcagaga	cttaggcagt	tcaacactta	aaaaattgca	7740
aacataacta	agaaaaacct	atgggagttt	gtttgctgca	gtttgctgca	cagaaaaata	7800
acatcagact	aacaacaagg	ttttgattag	ttttgcctgt	tagctgtatg	aagaaaaacat	7860
ttctaaatac	aaagttttct	aggaagttga	gtcactattt	aaaatccatt	gctttgacca	7920
atgtgaagcc	aacaggggtt	tcattcatat	agaatgcctt	tttatcaaaa	gagggtctga	7980
gtgttatacc	aaaatcagca	cccagctggt	ttcagtatgc	gcaagaaaac	atgaacgtcc	8040
cccaaaagcaa	atgtactcca	ttccgtgtta	accttatttc	tgctccctgg	tctctgatta	8100

```

gtctgggtga tgaatgtagc ttcccccttg gccatgcact cacagctttg aaggaggctt 8160
ctgaaggaac tggtagggaa ccctaggggc tcttgcta atgtgttgat agtcactgaa 8220
atcagtgatc accatgtcct ttcaactttga aaattctatg attctgtata atcaaacaac 8280
ctaactcact ggattttttca cttttcaacc agctattccg tgtcttacag tgctctgggc 8340
ttcacagata ggtaggcacc tggcatgtca gttgcctgaa tttgaaagt ttcacctgta 8400
tgttttggta cgataaaaaat aaaaatgtaa tttatatatc tgaatcaggt ctgtatgta 8460
tgatcaattg ctcagcaatt tcgggcaggt ggtttgatgg ttatgtagta atgtagcctg 8520
agagcagaaa tacagagcct ctgggctaga gaaagtataa atggcatcct aggcctatgta 8580
gggttacagc tcttcagaag gaactttcat tttcattgtg acacatcgtc tacatgttgt 8640
agaagaacat agtttcagaa ttcttccagt tagaaacata gtttctcaa tattcacttt 8700
cagcattggg tagaaaaagt acccagggtg tggactgatt gaggatccac tagtaacggc 8760
cgccagtgtg ctggaattct gcagatatcc atcacactgg cgcccgctcg agtctagagg 8820
gcccgtttaa acccgctgat cagcctcgac tgtgccttct agttgccagc catctgttgt 8880
ttgcccctcc cccgtgcctt ccttgaccct ggaagggtgc actcccactg tcctttccta 8940
ataaaatgag gaaattgcat cgcattgtct g 8971

```

<210> 43

<211> 29

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic primer

<400> 43

cgggtccccc gcgatacagt caactgtct

29

<210> 44

<211> 30

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic primer

<400> 44

taaagcttgc caccatgaag ctactgtctt

30

<210> 45

<211> 10514

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic recombinant  
DNA vector sequence

<400> 45

```

ctaaattgta agcgtaata ttttgtaaa attcggtta aatttttgtt aaatcagctc 60
attttttaac caataggccg aaatcggcaa aatcccttat aaatcaaaag aatagaccga 120
gatagggttg agtggtgttc cagtttgga caagagtcca ctattaaaga acgtggactc 180
caacgtcaaa gggcgaaaaa cgtctatca gggcgatggc ccactacgtg aaccatcacc 240
ctaataaagt tttttggggt cgagggtgcc taaagcacta aatcggaacc ctaaaggag 300
ccccgattt agagcttgac ggggaaagcc gggaacgtg gcgagaaagg aagggaagaa 360
agcgaaagga gcgggcgcta gggcgctggc aagtgtagcg gtcacgctgc gcgtaaccac 420

```

cacaccgcc	gcgcttaatg	cgccgctaca	ggcgcgctcc	cattcgccat	tcaggctgcy	480
caactgttgg	gaagggcgat	cgggtgcgggc	ctcttcgcta	ttacgccagc	tggcgaaaagg	540
gggatgtgct	gcaaggcgat	taagtgggt	aacgccaggg	tttcccagt	cacgacgttg	600
taaaacgacg	gccagtgcgc	gcgcgtaata	cgactcacta	tagggcgaat	tggagctcca	660
ccgcggtgcy	gccggggccat	gcaggccaga	catgataaga	tacattgatg	agtttggaca	720
aaccacaact	agaatgcagt	gaaaaaaatg	ctttatattgt	gaaatttgtg	atgctattgc	780
tttatttgta	accattataa	gctgcaataa	acaagttgtc	gagttacttg	tacagctcgt	840
ccatgccgag	agtgatcccg	gcgggcggtca	cgaactccag	caggaccatg	tgatcgcgct	900
tctcgttggg	gtctttgctc	agggcgggact	ggtagctcag	gtagtgggtg	tcgggcagca	960
gcacggggcc	gtcgccgatg	ggggtgttct	gctggtagtg	gtcggcgagc	tgacagctgc	1020
cgctctcgat	gttggtggcgg	atcttgaagt	tcaccttgat	gccgttcttc	tgcttgctcg	1080
ccatgatata	gacgttggtg	ctggtgtagt	tgtactccag	cttggtgcccc	aggatgttgc	1140
cgctctcctt	gaagtcgatg	cccttcagct	cgatgcgggt	caccagggtg	tcgccctcga	1200
acttcacctc	ggcgcggggtc	ttgtagttgc	cgctcgtcct	gaagaagatg	gtgcgctcct	1260
ggacgtagcc	ttcgggcatg	gcggacttga	agaagtcgtg	ctgcttcatg	tggtcggggt	1320
agcgggcgaa	gcactgcagg	ccgtagccga	aggtggtcac	gaggggtggg	cagggcacgg	1380
gcagcttgcc	ggtgggtgcag	atgaacttca	gggtcagctt	gccgtagggtg	gcacgcccct	1440
cgcctcgcc	ggacacgctg	aacttggtgc	cgtttacgtc	gccgtccagc	tcgaccagga	1500
tgggcaccac	cccggtgaac	agctcctcgc	ccttgctcac	catggtggcg	ttagcttgat	1560
ttgacagtgg	ctgggggttg	cgccgcccggg	ttttatagga	agccacagcg	gccactcgag	1620
ccataaaaagg	caactttagg	aacggcgggg	ggtgattgga	ttcgagtcgt	ttattcaccg	1680
gccttgccgc	acagtgcagc	atttttttgc	ccccctccc	ctccttttgc	gggggaaaaa	1740
aaaaaaaaaa	aaaaaaaaaag	gagaagagaa	aaaaagcgag	cgagagagaa	agcgagattg	1800
aggaagagga	tgaagagttt	ggcgatgggt	gctgggtccg	tagggccaga	tggacaagaa	1860
tagccccgc	ccttgccgac	agtatcccat	tcagtgactc	agatcagtc	agcggccgcg	1920
tctaacctag	gagcccacca	cagacactat	actataacca	cactgaattt	tctctcctat	1980
ctcaataat	cctattcttg	ttcatgcctc	agtgccttat	aacaagcaga	tttctcggcc	2040
agacatggtg	gcttgccgct	gtaatcccat	cactttggcc	aagtcgggca	gctcacttga	2100
ggtcaggagt	ttgaattcag	ctgttccaat	atgggtgaaac	cccatctcta	ctaaaaatac	2160
aaaaattagc	tgggcatggt	ggtgtgtgcc	tgtagtccca	gctactcagg	agactgaggc	2220
aggagaattg	cttgaactcg	ggaagcagag	gttgacgtga	gccaagatca	tgccattgca	2280
ctccagcctg	ggcaacagag	cgagactccg	tctcaaacaa	aacaaaacaa	aaacaaaaaa	2340
acaagcagat	ctctcaactt	ggaatgctgc	tgcttctcca	tgccaactgg	aaaacttctt	2400
tttcttttag	actcaactca	aatatgatcc	tcctgaagt	ctctccatct	ttcctagaca	2460
caattagcaa	ccagttcctc	tgtgttcaaa	cagcacttcg	tgtgtatgtg	tatgtatgtg	2520
tatttatcat	tattattatc	acctttgatc	actctccctg	aaggtaatag	ccaggcactt	2580
agtatagagt	aggatcaaac	aattgcttat	tgaaggata	aatgggtgaa	tgacaacttt	2640
gcaaagtttc	aatttttaat	ctattgattt	gaaagaatta	ttatacatta	ttttaaatat	2700
cctttctagt	tctaatatct	tatgtctata	gtagttcgac	agttacttta	aatcagatat	2760
taaatgattt	tcatttttaca	aagaacagac	ctgcctttct	caatgtcaca	agactattgt	2820
agagaatagc	agtgaataaa	gtacatgtga	aaataacttg	caaagttaata	gtgaaaatag	2880
agctgatctc	acaggaatgt	tttgaaaatt	aaaacaccgt	atgtaaaata	ggtaatctaa	2940
tctgtgggca	catattttaa	ctctcaataa	atattaactt	tgattattta	aaaaatacct	3000
ggtagctttt	ttgtgaattc	tacaagaacc	tgtacatgat	tggttgccat	ttccgtcaaa	3060
atgagaaaaat	tttcttctgc	actgaattct	tcttttaact	aaattttaag	aaaaaaacca	3120
atgattaaaa	agtagataag	aaaatgataa	tataaaacca	ttagactaaa	gaaaaactag	3180
gtttagatga	taaaatttga	aaacccatct	tactatttct	cagatctttg	gagggcaatt	3240
accattatca	agatagggtg	gagacttcat	actcttgaat	tgtttaatgt	tgctgagagt	3300
ttaaaagtct	aacttgaatt	ctcagctcta	atcatttcaa	agtaacatat	tcacatataa	3360
attatacctc	caaggcaata	aaaactccag	aaaatattaa	ctttaaacct	acataatttc	3420
agatattata	catacaattt	tatttggtat	ttcctgaggc	atcctctgtt	tgttatatga	3480
atccataata	aaatgtagaa	gagtcgtgtg	atctgggggtg	agttcagttt	tctcctacac	3540
tggcaagaaa	ataaaagggtg	ttagggaaga	tgtgtataaa	gccttctttc	tatatgtgca	3600
ttcatctgga	gggacctaat	agcagaaatc	ataaaaaaaa	aaaggcataa	atacagcagg	3660
aactatgagt	tctaagagag	tgtataattt	ctttctctgg	gtcctcaggc	aggccatatt	3720
acatttcact	gaattttcct	atctataaaa	tatgaataac	aaaatattcc	acttgtgtat	3780
ttcttaggaa	taaatggtag	tctggatatg	aagactattg	agccatttca	tattgaagggt	3840
cactgaaagt	tcaagcttta	attataacta	atgaaagggt	ttcatatagt	atgactacta	3900

tgaatactat	attcatatac	tattctacat	gaattccacc	tgggtgaagtc	atctgagagt	3960
attcatgtgg	aggagggatc	aacaaactgt	ggcccatgga	ccaaatccat	tggattttat	4020
aaagttttat	tggaacacag	tcacactcat	ttgtttatgt	attgtctctg	ctacttttgc	4080
actaaaatgg	cagagttgag	tagctgttgt	agataccatg	tggccaaagc	ccttttaaaga	4140
aaaaaagttt	gtcaacccct	aaaatagaga	gtcaacaatc	tggacaattc	agaccccaaa	4200
catctctttc	tccctctaca	gcactctttg	aagagtgtga	ataattatgt	caactctttg	4260
ttttcttata	ccaaagtgtt	actgatgaag	cactatggca	aaataatgct	ttcatataaa	4320
atactttttc	ttaaataaat	acaaaaatac	gtatttgtaa	aactaaatgt	ggataaaagt	4380
ttctttgtca	ataatgtaat	aaatattgag	aaggggtgtg	tttggacaat	catagcttac	4440
tggtttttca	gctaattggc	aaataaaatt	ccatcaaattg	tcctgtgcat	tatcacaatg	4500
tgtgtgtcta	aacaccttgg	atgctgttcc	cctattttgt	cctaagttgc	tcaatggatt	4560
tcttgagaaa	tttghtaacc	cagtgggtag	gactttttct	tattcatcca	catacgtgta	4620
ttcacctgt	tttactcatt	tataaggatt	aatccacagc	ctctattgag	accatgttct	4680
tcccagtttg	agactgtctt	ctcaaaaaga	gggaaaaaat	ggtcatcaat	actcagttgg	4740
aaaaaatcca	ttttatacac	caaatacatt	ctgctggcaa	acactaggtg	gtaggactgg	4800
aatggagtc	gagagaagga	aaactttcac	tttttagttt	atgcattgct	actatttgaa	4860
tgtttacatc	aaagcatgta	ttaggaatcc	attttaatga	acaattggac	tagaaattca	4920
gctgtgcctg	tggctttaat	ttgaaagcct	cctggccagc	tcactatgta	cctagaaaact	4980
cagttcctgc	cagtcttggc	agacaccatt	gcatattatt	accctgcatg	actttgttgt	5040
cgaggtcact	tgtcgcaagt	cacgaccttc	actgtcttct	ctagacggcg	atacagtcaa	5100
ctgtctttga	cctttgttac	tactctcttc	cgatgatgat	gtcgcactta	ttctatgctg	5160
tctcaatgtt	agaggcatat	cagttctccac	tgaagccaat	ctatctgtga	cggcatcttt	5220
attcacatta	tcttgtacaa	ataatcctgt	taacaatgct	tttatatcct	gtaaaagaatc	5280
cattttcaaa	atcatgtcaa	ggtcttctcg	aggaaaaatc	agtagaaata	gctgttccag	5340
tctttctagc	cttgattcca	cttctgtcag	atgtgcccta	gtcagcggag	accttttggt	5400
tttgggagag	tagcgacact	cccagttggt	cttcagacac	ttggcgcaact	tcggtttttc	5460
tttgaggcac	ttgagctttt	taagtccggc	aatatcgcac	gcttggttca	tagaagacag	5520
tagcttctgc	gtggcggatc	cgataagcc	agtaagcagt	gggttctcta	gttagccaga	5580
gagctctgct	tatatagacc	tcccacgta	cagccctacc	gccatttgc	gtcaatgggg	5640
cggagtgtgt	acgacatttt	ggaaagtccc	gttgattttg	gtgccaaaac	aaactcccat	5700
tgacgtcaat	ggggtggaga	cttggaatc	cccgtgagtc	aaaccgctat	ccacgcccac	5760
tgatgtactg	ccaaaaccgc	atcaccatgg	taatagcgat	gactaatacg	tagatgtact	5820
gccaaagttag	aaagtcccat	aaggtcatgt	actgggcata	atgccaggcg	ggccattttac	5880
cgtcattgac	gtcaataggg	ggcgtacttg	gcatttgata	caattgatgt	actgccaagt	5940
gggcagttta	ccgtaaatac	tccaccatt	gacgtcaatg	gaaagtccct	attggcggtta	6000
ctatgggaac	atacgtcatt	attgaagtc	atgggcgggg	gtcgttgggc	ggtcagccag	6060
gcgggccatt	taccgtaagt	tatgtaacgc	ggaactccat	atatgggcta	tgaactaatg	6120
accccgtaat	tgattactat	taataactag	tcaataatca	atgtcaacgc	gtatatctgg	6180
cccgtacatc	gcgaagcagc	gcaaaacgcc	taaccctacc	cgggctgcag	gaattctacc	6240
gggtagggga	ggcgcttttc	ccaaggcagt	ctggagcatg	cgcttttagca	gccccgctgg	6300
cacttggcgc	atcacaagtg	gcctctggcc	tcgcacacat	tccacatcca	ccggtagcgc	6360
caaccggctc	ccttcttttg	tggccccttc	gcgccacctt	ctactcctcc	cctagtcagg	6420
aagtcccccc	ccgccccgca	gctcgcgtcg	tgcaggacgt	gacaaatgga	agtagcacgt	6480
ctcactagtc	tcgtgcagat	ggacaagcac	cgctgagcaa	tggaaagcggg	taggcctttg	6540
gggcagcggc	caatagcagc	ttggctcctt	cgctttctgg	gctcagaggc	tgggaagggg	6600
tgggtccggg	ggcgggctca	ggggcgggct	caggggcggg	gcgggcgcga	aggtcctccg	6660
gaccggcat	tctgcacgct	tcaaaagcgc	acgtctgccg	cgctgttctc	ctcttctcca	6720
tctccgggcc	ttcgacctgc	atgaaaaagc	ctgaactcac	cgcgacgtct	gtcgagaagt	6780
ttctgatcga	aaagttcgac	agcgtctccg	acctgatgca	gctctcggag	ggcgaagaat	6840
ctcgtgcttt	cagcttcgat	gtaggagggc	gtggatatgt	cctgcgggta	aatagctgcg	6900
ccgatggttt	ctacaaagat	cgttatgttt	atcggcactt	tgcacggcc	gcgctcccga	6960
ttccggaagt	gcttgacatt	ggggaattca	gcgagagcct	gacctattgc	atctcccgcc	7020
gtgcacagg	tgtcacgttg	caagacctgc	ctgaaaccga	actgcccgt	gttctgcagc	7080
cggctcgcga	ggccatggat	gcgatcgctg	cggccgatct	tagccagacg	agcgggttcg	7140
gcccattcgg	accgcaagga	atcgggtcaat	acactacatg	gcgtgatttc	atatgcgcga	7200
ttgctgatcc	ccatgtgtat	cactggcaaa	ctgtgatgga	cgacaccgtc	agtgcgtccg	7260
tcgcgcaggc	tctcgatgag	ctgatgcttt	gggcccagga	ctgccccgaa	gtccggcacc	7320
tcgtgcacgc	ggatttcggc	tccaacaatg	tcctgacgga	caatggccgc	ataacagcgg	7380



tcattgactg	gagcgaggcg	atgttcgggg	attcccaata	cgaggtcgcc	aacatcttct	7440
tctggaggcc	gtgggtggct	tgtatggagc	agcagacgcg	ctacttcgag	cggaggcatc	7500
cggagcttgc	aggatcgccg	cggctccggg	cgtatatgct	ccgcattggt	cttgaccaac	7560
tctatcagag	cttgggttgac	ggcaatttcg	atgatgcagc	ttggggcgag	ggtcgatgcg	7620
acgcaatcgt	ccgatccgga	gccggggactg	tcgggcgtac	acaaatcgcc	cgcagaagcg	7680
cggccgtctg	gaccgatggc	tgtgtagaag	tactcgccga	tagtggaac	cgacgccccca	7740
gcactcgtcc	gagggcacaag	gaatagagta	tgcagaaatt	gatgatctat	taaaacaataa	7800
agatgtccac	taaaatggaa	gtttttcctg	tcatactttg	ttaagaaggg	tgagaacaga	7860
gtacctacat	tttgaatgga	aggattggag	ctacgggggt	gggggtgggg	tgggattaga	7920
taaatgcctg	ctctttactg	aaggctcttt	actattgctt	tatgataatg	tttcatagtt	7980
ggatatcata	attttaaaca	gcaaaaacca	attaagggcc	agctcattcc	ttccactcat	8040
gatctataga	tctatagatc	tctcgtggga	tcattgtttt	tctcttgatt	cccactttgt	8100
ggttctaagt	actgtggttt	ccaaatgtgt	cagtttcata	gcctgaagaa	cgagatcagc	8160
agcctctgtt	ccacatacac	ttcattctca	gtattgtttt	gccaaagtct	aattccatca	8220
gatttttagtg	gacatcttta	ttgtttaata	gatcatcaat	ttctgcatac	taagcttatc	8280
gataccgtcg	acctcgaggg	ggggcccgtt	acccagcttt	tgttcccttt	agtgagggtt	8340
aattgcgcgc	ttggcgtaat	catggtcata	gctgtttcct	gtgtgaaatt	gttaccgct	8400
cacaattcca	cacaacatac	gagccggaag	cataaagtgt	aaagcctggg	gtgcctaattg	8460
agtgagctaa	ctcacattaa	ttgcgttgcg	ctcactgccc	gctttccagt	cgggaaacct	8520
gtcgtgccag	ctgcattaat	gaatcggcca	acgcgcgggg	agaggcggtt	tgcgtattgg	8580
gcgctcttcc	gcttcctcgc	tcaactgactc	gctgcgctcg	gtcgttcggc	tgcggcgagc	8640
ggatcagct	cactcaaagg	cggtaataacg	gttatccaca	gaatcagggg	ataacgcagg	8700
aaagaacatg	tgagcaaaaag	gccagcaaaa	ggccagggaac	cgtaaaaagg	ccgcgttgct	8760
ggcgtttttc	cataggctcc	gccccctga	cgagcatcac	aaaaatcgac	gctcaagtca	8820
gaggtggcga	aaccgcagac	gactataaag	ataccaggcg	tttccccctg	gaagctccct	8880
cgtgcgctct	cctgttccga	ccctgcccgt	taccggatac	ctgtccgcct	ttctcccttc	8940
gggaagcgtg	gcgctttctc	atagctcaag	ctgtaggtat	ctcagttcgg	tgtaggtcgt	9000
tcgctccaag	ctgggctgtg	tgcacgaacc	ccccgttcag	cccagaccgt	gcgccttatc	9060
cggtaactat	cgtcttgagt	ccaaccgggt	aagacacgac	ttatcgccac	tggcagcagc	9120
cactggtaac	aggattagca	gagcgaggta	tgtaggcggt	gctacagagt	tcttgaagtg	9180
gtggcctaac	tacggctaca	ctagaaggac	agtatttggt	atctgcgctc	tgtgaagcc	9240
agttaccttc	ggaaaaagag	ttggtagctc	ttgatccggc	aaacaaacca	ccgctggtag	9300
cgggtggtttt	tttgtttgca	agcagcagat	tacgcgcaga	aaaaaaggat	ctcaagaaga	9360
tcctttgatc	ttttctacgg	ggcttgacgc	tcagtggaac	gaaaactcac	gttaagggat	9420
tttggtcattg	agattatcaa	aaaggatcct	caccatagatc	cttttaaatt	aaaaatgaag	9480
ttttaaatca	atctaaagta	tatatgagta	aacttggtct	gacagttacc	aatgcttaat	9540
cagtgaggca	cctatctcag	cgatctgtct	atttcgttca	tccatagttg	cctgactccc	9600
cgtcgtgtag	ataactacga	tacgggaggg	cttaccatct	ggccccagtg	ctgcaatgat	9660
accgcgagac	ccacgctcac	cggctccaga	tttatcagca	ataaaccagc	cagccggaag	9720
ggccgagcgc	agaagtggtc	ctgcaacttt	atccgcctcc	atccagtcta	ttaattgttg	9780
ccgggaagct	agagtaagta	gttcgccagt	taatagtttg	cgcaacgttg	ttgccattgc	9840
tacaggcatc	gtgggtgtcac	gctcgtcggt	tggtaggct	tcattcagct	ccggttccca	9900
acgatcaagg	cgagttacat	gatcccccat	gttgtgcaaa	aaagcggtta	gctccttcgg	9960
tcctccgatc	gttgtcagaa	gtaagttggc	cgcagtgtta	tcactcatgg	ttatggcagc	10020
actgcataat	tctcttactg	tcatgccatc	cgtaagatgc	ttttctgtga	ctggtgagta	10080
ctcaaccaag	tcattctgag	aatagtgtat	gcggcgaccg	agttgctctt	gcccggcgctc	10140
aatacgggat	aataccgcgc	cacatagcag	aactttaaaa	gtgctcatca	ttggaaaaacg	10200
ttcttcgggg	cgaaaactct	caaggatctt	accgctgttg	agatccagtt	cgatgtaacc	10260
cactcgtgca	cccaactgat	cttcagcatc	ttttactttc	accagcgttt	ctgggtgagc	10320
aaaaacagga	aggcaaaatg	ccgcaaaaaa	gggaataagg	gcgacacgga	aatgttgaat	10380
actcatactc	ttcctttttc	aatattattg	aagcatttat	cagggttatt	gtctcatgag	10440
cggatacata	tttgaatgta	tttagaaaaa	taaaacaata	ggggttccgc	gcacatttcc	10500
ccgaaaagtg	ccac					10514

<210> 46  
 <211> 9626  
 <212> DNA

## &lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Synthetic recombinant DNA vector sequence

&lt;400&gt; 46

gacgcatcgg	gagatctccc	gatccccctat	ggctgactct	cagtacaatc	tgctctgatg	60
ccgcatagtt	aagccagtat	ctgctccctg	cttggtgtgt	ggaggctcgt	gagtagtgcg	120
cgagcaaaat	ttaagctaca	acaaggcaag	gcttgaccga	caattgcatg	aagaatctgc	180
ttaggggttag	gcgtttttg	ctgcttcg	atgtacgggc	cagatatacg	cggtgacatt	240
gattattgac	tagttattaa	tagtaataca	ttacgggggtc	attagttcat	agcccatata	300
tggagttccg	cgttacataa	cttacggtaa	atggcccgc	tggctgaccg	cccaacgacc	360
cccgcacatt	gacgtcaata	atgacgtatg	ttcccatagt	aacgcccaata	gggactttcc	420
attgacgtca	atgggtggac	tatttacggg	aaactgcccc	cttggcagta	catcaagtgt	480
atcatatgcc	aagtacgccc	cctattgacg	tcaatgacgg	taaatggccc	gcctggcatt	540
atgccagta	catgacctta	tgggactttc	ctacttggca	gtacatctac	gtattagtca	600
tcgctattac	catggtgatg	cggttttggc	agtacatcaa	tgggcgtgga	tagcggtttg	660
actcacgggg	atttccaagt	ctccacccca	ttgacgtcaa	tgggagtttg	ttttggcacc	720
aaaatcaacg	ggactttcca	aaatgtcgta	acaactccgc	cccattgacg	caaatggg	780
gtaggcgtgt	acgggtgggag	gtctatataa	gcagagctct	ctggctaact	agagaaccca	840
ctgcttactg	gcttatcgaa	attaatacga	ctcactatag	ggagacccaa	gctggctagc	900
gtttaaactt	aagcttggta	ccgagctcgg	atccactagt	ccagtgtggt	ggaattctgc	960
agatatccag	cacagtggct	tgcggccg	aaggaagata	gccaaggctt	actgaggctg	1020
gtggaggagg	ccactgctgg	gtcaccatg	gaccgcccga	tgtggggggc	ccacgtcttc	1080
tgcgtgttga	gcccgttacc	gaccgtgagt	agccagctga	gaccctctgg	cctctgccac	1140
tgggctccag	atttgggagc	cacaggggca	actggagcct	ggcggagacc	ctactgccct	1200
tctcctgctc	tagagtcctc	ccctacgagc	agggtggcag	ctgtggcttg	gagagccctc	1260
gacactgggg	ctccctccct	catctccctc	tgcccagcaa	tctgtagtga	ctctgggtag	1320
tgctattaac	atgtccttcc	aggggctgag	ctgagctcac	cttctgtgtg	cttccccccag	1380
ggcctggctg	tgctctgagg	ggccacagag	actctgtctg	ccacaagaca	gtgcctcaga	1440
attggtgagg	caaagtcagg	ctccctctct	agtccttgca	gtttgtctct	tctctcagct	1500
gtgcacctgt	gatgcttcca	cacgtcctct	tggttaagctt	tctggaccct	tccttagccc	1560
tatgactccc	tgagccaact	tccagttttg	tggtttcctt	tgaagagggt	gtggccaggg	1620
aaggaaatgga	acctaaggag	ttggtggaga	ccagctgctt	ctgctccagc	acttagggga	1680
tccatcatct	tctcacgatg	ctgagggtcg	gagaggcttt	tggcaaccaa	attgtactcc	1740
atgggctaga	gcgctggagc	acggggcatc	aggagagagg	gcccttctct	aatgggctgc	1800
agtttcccag	taactttcca	ggaggttact	gcaagtgaca	gcaggggacca	ggctgggagt	1860
cacaaggttc	tgggatgggtg	agggcctgaa	gcattcttcc	tccgctgcct	tattggaggg	1920
tgctgcccct	cccgagcacc	ttgctgcacc	tgctcctgag	tgacagggct	ttgagcttga	1980
atcctggacg	catgaaacta	gccgtgtgaa	ctgcaagttc	ctcagccttc	ctgagctctg	2040
cttttatctc	ttttgttaaa	agagaaaata	tccgctggcc	cgggtcactg	tgcatgtagt	2100
ggaatctgat	ttgtcaaggg	ctgggcatag	gggtgcaaca	caggaaactt	cgactgcccc	2160
cctgtctgcc	atctgtgtgc	actctacagc	aaaagcttta	tttttagcag	gggtacccaa	2220
gatctggtgg	tggccattca	gcagatgggtg	ctagtaccaa	agctagaact	caatcactgt	2280
ttgccagtgt	gggcatcatc	ccctgcagat	accaggaaga	gcagcaggca	tagatgcaag	2340
gcctgggggt	ccagggcgat	ggagcctccc	ctcaggggtg	ggtgggggtcc	acgatgtgtg	2400
tgggggggtg	actgggagcc	ccaggatgtg	tttcgccagt	cctaattcag	caggagcatt	2460
tggataaatc	ttagaacgat	ctatccttcc	actcccacca	cttccctccc	tagaatctct	2520
ccttacatct	ggccctatgc	ccacacatcc	tgagtcatta	gtggggcttc	tctgaatttc	2580
cctccttctc	ttagagggtgt	gggcactgtc	tgttcatcct	tttggcatga	ggctgctctc	2640
ctggcatcct	ttgccctcag	ttgcttcccc	aaattggccc	tgaagagttt	cccaggcagg	2700
gttgggggtt	gccagactgt	gcagtaggca	gaaggatagg	cagagaccac	agctcttcca	2760
acagcccagg	gtcaatatca	ggctgatgggt	gtgagaaagt	cccatacagg	cccatctgcc	2820
ctcctgctct	tccttccccg	ctttgtgcgc	tcagtctgtt	gggagatgta	ggggagacgc	2880
aacgggggtc	tgccctgaga	tggacctgag	gttactgagt	cctcgggagg	aagccaagag	2940
ctgagtcagg	ggcaccaccc	aaaataagac	ctggagcctg	ggtagtgggt	gacatctgca	3000
ccctggggcc	ttgcaggggc	aggcaggaga	tggcagagct	gcgtccaagc	catggaccca	3060

tgggggtcagc	agggcggggac	aggagtgact	tctcctgcc	gggtagcact	ggtgggggtg	3120
accctcctta	ctggggtggc	tctccagatt	ttctcctatc	ccacaaatct	gcggcagagc	3180
tctcgggtct	ctccacgtgg	tcacccacca	tctattgtcc	ccactctggt	ggtcaaagtc	3240
tctactctct	gaccccagcc	ctggtcttac	atctgctaca	ctatttgctg	tgctgcatt	3300
tctcccaaaag	aacactagct	gttaattaaa	catttgctga	catggatcct	gctagacatt	3360
gaggggtgtgt	gtgtgtgggg	tgggggtggg	ggccatgggg	agagaaacat	tagtcctgag	3420
ctgatacttc	ccaaatacac	cttttcccac	ttcccttgct	cttgcacatg	ctgtgctttc	3480
tgttccagat	gctctttctcc	ccactgtgtg	tgtctggctc	taaaaaatgc	tccttctgtcc	3540
ccagggtccag	ctcaaagtgc	acctccccaa	tcctctgtgag	agggacattc	cctcccatgg	3600
tacctgtctt	gcttcttttt	gttttttttt	tgtttttctg	agatggagtc	tcactttatc	3660
gcccaggctg	gcagacagtg	gcacaatctc	agctcactgt	aacctccgcc	tcctgggttc	3720
aagtgattct	cctgcctcag	cctcctgagt	agctgggatt	acaggagcgt	gtcatcacac	3780
ccagctaatt	tttgattttt	tgtagagata	gtgttttgcc	acattggtca	ggctggtgtt	3840
taactcctga	cgtcagggtga	tccgcctacc	ttggcctccc	aaagtgtctg	gattacaggc	3900
atgagccact	gcgcccggcc	cctgtcctgc	ttctgatcac	agtttgcgca	gggtctatta	3960
ggatccagcc	cttctccctt	ccccagcccc	tcctctcctg	tgaagcctg	ggggaacaga	4020
catctacact	ggttgacttg	gatgtgtgtt	ttatacaaaa	atgtgtatgc	tgtataccca	4080
gttcacagtg	cagaaagtat	gttttgtaaa	cttggcttcc	cttgtatgta	tctgtttggg	4140
ctctggcctg	ctgggcgatc	agttttgggt	ttccttcggg	ttggggcggt	gggagggaga	4200
agtgtctagaa	attcctgctg	ttcacttcca	gtttcctgga	gaacttggtt	aaatacaaaag	4260
tgctggatcc	accctagatc	tctacatatt	cctgcttcca	tcccaatgga	tctctagctg	4320
gttcttgagg	cacactttca	gaagtactgc	tagaagattc	ccaaaaaaac	cccaacttgt	4380
ttccagcttc	cagagatcat	ttctcatggc	cactgtgtag	cacaagatct	gggaggctcc	4440
taaaactata	atctttcaag	gttgctcagc	catcaagggt	atgatccttt	catctacaca	4500
tccatgaagc	cacacatcca	tccttacaca	gctctctcac	tcatccacca	tgaccccatc	4560
catgcagcca	gtcatctgtc	tatccatcca	tccacctccc	cagccatata	tttgtctgtc	4620
tgtctgtata	ctcatccatt	ccactagcct	tccttccttc	ttttctgcct	tcctgccttc	4680
ctgccttcct	tccttccttc	cttttttttt	gcagccttca	cctcattcat	gtgtcactc	4740
atccatgtgc	tcatttcatt	atccatcttc	tctctcactt	actaagctac	ccatttgctc	4800
atgtgttcct	gcctgtgtgc	acccctaaaa	gcagagactg	agacaaaggt	cagaaattta	4860
aaatggtttc	cagtggggta	aatccaaagt	attggcagag	ctgcattcct	cctggagact	4920
gtagagaata	atccattttc	ttgccttttc	cttgggttca	tgccctttct	ccatcttcag	4980
accagcagtg	tggcatcttt	aagcggccgc	tcgagtctag	agggcccgtt	taaacccgct	5040
gatcagctc	gactgtgcct	tctagtggcc	agccatctgt	tgtttgcccc	tccccctgct	5100
cttccttgac	cctggaaggt	gccactccca	ctgtcctttc	ctaataaaaat	gaggaaattg	5160
catcgcatgt	tctgagtagg	tgtcatttca	ttctgggggg	tgggggtggg	caggacagca	5220
agggggagga	ttgggaagac	aatagcaggc	atgctgggga	tgcggtgggc	tctatggctt	5280
ctgaggcgga	aagaaccagc	tgggggtcta	gggggtatcc	ccacgcgccc	tgtagcggcg	5340
cattaagcgc	ggcgggtgtg	gtgggtacgc	gcagcgtgac	cgctacactt	gccagcggcc	5400
tagcggccgc	tcctttcgct	ttcttccctt	cctttctcgc	cacgttcgcc	ggctttcccc	5460
gtcaagctct	aaatcggggc	atccctttag	ggttccgatt	tagtgcttta	cggcacctcg	5520
acccccaaaa	acttgattag	gggtgatggt	cacgtagtgg	gccatcgccc	tgatagacgg	5580
tttttcgccc	tttgacgttg	gagtccacgt	tctttaatag	tggactcttg	ttccaaactg	5640
gaacaacact	caaccctatc	tcggtctatt	cttttgattt	ataagggtat	ttggggattt	5700
cggcctattg	gttaaaaaat	gagctgattt	aacaaaaatt	taacgcgaat	taattctgtg	5760
gaatgtgtgt	cagttagggt	gtggaaagtc	cccaggctcc	ccaggcaggc	agaagtatgc	5820
aaagcatgca	tctcaattag	tcagcaacca	gggtgtgaaa	gtcccaggc	tcccagcag	5880
gcagaagtat	gcaaagcatg	catctcaatt	agtcagcaac	catagtcccg	cccctaactc	5940
cgccccatccc	gcccctaact	ccgcccagtt	ccgcccattc	tccgccccat	ggctgactaa	6000
tfttttttat	ttatgcagag	gccgaggccg	cctctgcctc	tgagctattc	cagaagtagt	6060
gaggaggctt	ttttggaggc	ctaggctttt	gcaaaaagct	cccgggagct	tgtatatcca	6120
ttttcggtac	tgatcagcac	gtgatgaaaa	agcctgaact	caccgcgacg	tctgtcgaag	6180
agttttctgat	cgaaaagtcc	gacagcgtct	ccgacctgat	gcagctctcg	gagggcgaaag	6240
aatctcgtgc	tttctagctt	gatgtaggag	ggcgtggata	tgtcctgcgg	gtaaaatagct	6300
gcgcccgatgg	tttctacaaa	gatcgttatg	tttatcgcca	ctttgcatcg	gccgcgctcc	6360
cgattccgga	agtgttgac	attgggggaat	tcagcgagag	cctgacctat	tgcattctcc	6420
gccgtgcaca	gggtgtcacg	ttgcaagacc	tgcttgaac	cgaactgcc	gctgttctgc	6480
agccggtcgc	ggaggccatg	gatgcgatcg	ctgcggccga	tcttagccag	acgagcgggt	6540

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Synthetic primer

&lt;400&gt; 47

gagaacccac tgcttactgg ct

22

&lt;210&gt; 48

&lt;211&gt; 21

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Synthetic primer

&lt;400&gt; 48

gcattgctgt gcagtctgta a

21

&lt;210&gt; 49

&lt;211&gt; 7354

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence: Synthetic recombinant  
DNA vector sequence

&lt;400&gt; 49

cgcgattgct	gatcccatg	tgtatcactg	gcaaactgtg	atggacgaca	ccgtcagtgc	60
gtccgtcgcg	caggctctcg	atgagctgat	gctttgggcc	gaggactgcc	ccgaagtccg	120
gcacctcgtg	cacgcggatt	tcggctccaa	caatgtcctg	acggacaatg	gccgcataac	180
agcggtcatt	gactggagcg	aggcgatggt	cggggattcc	caatacgagg	tcgccaacat	240
cttcttctgg	aggccgtggt	tggcttgat	ggagcagcag	acgcgctact	tcgagcggag	300
gcatccggag	cttgaggat	cgccgcggct	ccgggcgtat	atgctccgca	ttggctctga	360
ccaactctat	cagagcttgg	ttgacggcaa	tttcgatgat	gcagcttggg	cgcagggtcg	420
atgcgacgca	atcgtccgat	ccggagccgg	gactgtcggg	cgtacacaaa	tcgcccgcag	480
aagcgcggcc	gtctggaccg	atggctgtgt	agaagtactc	gccgatagtg	gaaaccgacg	540
ccccagcact	cgtccgaggg	caaaggaata	gagtatgcag	aaattgatga	tctattaaac	600
aataaagatg	tccactaaaa	tgggaagttt	tccgtgcata	ctttgttaag	aagggtgaga	660
acagagtacc	tacattttga	atggaaggat	tggagctacg	gggggtgggg	tgggggtggga	720
ttagataaat	gcctgctctt	tactgaaggc	tctttactat	tgctttatga	taatgtttca	780
tagttggata	tcataattta	aacaagcaaa	accaaattaa	gggccagctc	attcctccca	840
ctcatgatct	atagatctat	agatctctcg	tgggatcatt	gtttttctct	tgattcccac	900
tttgtgggtc	taagtactgt	ggtttccaaa	tgtgtcagtt	tcatagcctg	agaacgaga	960
tcagcagcct	ctgttccaca	tacacttcat	tctcagtatt	gttttgccaa	gttctaattc	1020
catcagattt	tagtggacat	ctttattggt	taatagatca	tcaatttctg	cataactaagc	1080
ttatcgatac	cgtcgacctc	gagggggggc	ccggtaccca	gcttttggtc	cctttagtga	1140
gggttaattg	cgcgcttggc	gtaatcatgg	tcatagctgt	ttcctgtgtg	aaattgttat	1200
ccgctcacaa	ttccacacaa	catacgagcc	ggaagcataa	agtgtaaagc	ctgggggtgcc	1260
taatgagtga	gctaactcac	attaattgcg	ttgcgctcac	tgcccgcctt	ccagtcggga	1320
aacctgtcgt	gccagctgca	ttaatgaatc	ggccaacgcg	cggggagagg	cggtttgcgt	1380
attgggcgct	cttcgcgttc	ctcgcctcact	gactcgctgc	gctcggctcg	tcggctgcgg	1440
cgagcgggtat	cagctcactc	aaaggcggta	atacggttat	ccacagaatc	aggggataac	1500
gcaggaaaga	acatgtgagc	aaaaggccag	caaaaggcca	ggaaccgtaa	aaaggccgcg	1560
ttgctggcgt	ttttccatag	gctccgcccc	cctgacgagc	atcacaaaaa	tcgacgctca	1620
agtcagaggt	ggcgaaaccc	gacaggacta	taaagatacc	aggcgtttcc	ccctggaagc	1680

tccctcgtgc	gctctcctgt	tccgaccctg	ccgcttacccg	gatacctgtc	cgcctttctc	1740
ccttcgggaa	gcgtggcgct	ttctcatagc	tcacgctgta	ggtatctcag	ttcgggtgtag	1800
gtcgttcgct	ccaagctggg	ctgtgtgcac	gaaccccccg	ttcagcccca	ccgctgcgcc	1860
ttatccggta	actatcgtct	tgagtccaac	ccggtaagac	acgacttatc	gccactggca	1920
gcagccactg	gtaacaggat	tagcagagcg	aggatgtag	gcgggtgctac	agagttcttg	1980
aagtgggtggc	ctaactacgg	ctacactaga	aggacagtat	ttgggtatctg	cgctctgctg	2040
aagccagtta	ccttcggaaa	aagagttggg	agctcttgat	ccggcaaaaa	aaccaccgct	2100
ggtagcgggtg	gtttttttgt	ttgcaagcag	cagattacgc	gcagaaaaaa	aggatctcaa	2160
gaagatcctt	tgatcttttc	tacgggggtct	gacgctcagt	ggaacgaaaa	ctcacgttaa	2220
gggatttttg	tcatagagatt	atcaaaaaag	atcttcacct	agatcctttt	aaattaaaaa	2280
tgaagtttta	aatcaatcta	aagtatatat	gagtaaactt	gggtctgacag	ttaccaatgc	2340
ttaatcagtg	aggcacctat	ctcagcgatc	tgtctatctt	gttcatccat	agttgcctga	2400
ctccccgtcg	tgtagataac	tacgatacgg	gagggcttac	catctggccc	cagtgtctga	2460
atgataccgc	gagaccacag	ctcaccggct	ccagatttat	cagcaataaa	ccagccagcc	2520
ggaagggccg	agcgcagaag	tggtcctgca	actttatccg	cctccatcca	gtctattaat	2580
tggtgccggg	aagctagagt	aagtagttcg	ccagttaata	gtttgcgcaa	cgttggtgcc	2640
attgctacag	gcatcgtggg	gtcacgctcg	tcgtttggta	tggcttcatt	cagctccggt	2700
tccaacgat	caaggcgagt	tacatgatcc	cccatgttgt	gcaaaaaagc	ggttagctcc	2760
ttcggtcctc	cgatcgttgt	cagaagtaag	ttggccgcag	tgttatcact	catgggtatg	2820
gcagcactgc	ataattctct	tactgtcatg	ccatccgtaa	gatgcttttc	tgtgactggg	2880
gagtactcaa	ccaagtcatt	ctgagaatag	tgtatgcggc	gaccgagttg	ctcttgcccc	2940
gcgtcaatac	gggataatac	cgcgccacat	agcagaactt	taaaagtgtc	catcattgga	3000
aaacgttctt	cggggcgaaa	actctcaagg	atcttaccgc	tggttgagatc	cagttcagtg	3060
taaccctac	gtgcacccaa	ctgatcttca	gcatctttta	ctttcaccag	cgtttctggg	3120
tgagcaaaaa	caggaaggca	aaatgccgca	aaaaagggaa	taagggcgac	acggaaatgt	3180
tgaatactca	tactcttcct	ttttcaatat	tattgaagca	tttatcaggg	ttattgtctc	3240
atgagcggat	acatatattga	atgtatttag	aaaaataaac	aaataggggt	tccgcgcaca	3300
tttccccgaa	aagtgccacc	taaattgtaa	gcgttaatat	tttgttaaaa	ttcgcgttaa	3360
atttttgtta	aatcagctca	tttttttaacc	aataggccga	aatcggcaaa	atcccttata	3420
aatcaaaaga	atagaccgag	atagggttga	gtgttggtcc	agtttggaac	aagagtccac	3480
tattaaagaa	cgtggactcc	aacgtcaaag	ggcgaaaaaac	cgtctatcag	ggcgatggcc	3540
cactacgtga	accatcaccc	taatcaagtt	ttttgggggc	gaggtgccgt	aaagcactaa	3600
atcggaaccc	taaagggagc	ccccgattta	gagcttgacg	gggaaagccg	gcgaacgtgg	3660
cgagaaagga	agggaaagaa	gcgaaaggag	cgggcgctag	ggcgctggca	agtgtagcgg	3720
tcacgctcg	cgtaaccacc	acaccgcggc	cgcttaatgc	gccgctacag	ggcgctccc	3780
attcgccatt	caggctgcgc	aactgttggg	aagggcgatc	gggtgcgggc	tcttcgctat	3840
tacgccagct	ggcgaaagg	ggatgtgctg	caaggcgatt	aagttgggta	acgccagggt	3900
tttcccagtc	acgacgttgt	aaaacgacgg	ccagtgcgag	cgcgtaatac	gactcactat	3960
agggcgaaat	ggagctccac	cgcgggtgcg	ccgggccatg	caggccagac	atgataagat	4020
acattgatga	gtttggacaa	accacaacta	gaatgcagt	aaaaaaatgc	tttatttgtg	4080
aaatttgtga	tgctattgct	ttatttgtaa	ccattataag	ctgcaataaa	caagtgtctg	4140
agttacttgt	acagctcgtc	catgccgaga	gtgatcccg	cggcgggtcac	gaactccagc	4200
aggaccatgt	gatcgcgctt	ctcgttgggg	tctttgctca	gggcggactg	gtagctcagg	4260
tagtggttgt	cgggcagcag	cacggggccg	tcgccgatgg	gggtgttctg	ctggtagtgg	4320
tcggcgagct	gcacgctgcc	gtcctcgatg	ttgtggcgga	tcttgaagtt	caccttgatg	4380
ccgttcttct	gcttgtcggc	catgatatag	acgttgtggc	tgttgtagtt	gtactccagc	4440
ttgtgcccc	ggatgttgcc	gtcctccttg	aagtcgatgc	ccttcagctc	gatgcggttc	4500
accaggggtg	cgcctcgaa	cttcacctcg	gcgcgggtct	tgtagttgcc	gtcgtccttg	4560
aagaagatgg	tgcgctcctg	gacgtagcct	tcgggcattg	cggacttgaa	gaagtcgtgc	4620
tgcttcatgt	ggtcggggta	gcggggcgaag	cactgcaggc	cgtagccgaa	ggtggtcacg	4680
aggggtggcc	agggcacggg	cagcttgccg	gtgggtgcaga	tgaacttcag	ggtcagcttg	4740
ccgtaggtgg	catcgccctc	gccctcgccg	gacacgctga	acttgtggcc	gtttacgtcg	4800
ccgtccagct	cgaccaggat	gggcaccacc	ccggtgaaca	gctcctcgcc	cttgctcacc	4860
atgggtggcgt	tagcttgatt	tgacagtggc	tgggggttgc	gccgcgggtt	tttataggaa	4920
gccacagcgg	ccactcgagc	cataaaaggc	aacttttagga	acggcggggg	gtgattggat	4980
tcgagtcgtt	tattcaccgg	ccttgccgca	cagtgcagca	tttttttacc	ccctctcccc	5040
tccttttgcg	ggggaaaaaa	aaaaaaaaaa	aaaaaaaaag	agaagagaaa	aaaagcgagc	5100
gagagagaaa	gcgagattga	ggaagaggat	gaagagtttg	gcgatgggtg	ctggttccgt	5160

```

aggccccagat ggacaagaat agcccccgcc cttgcggaaca gtatcccatt cagtgactca 5220
gatcagtc aa ggcggcgctc tagacggcga tacagtcaac tgtctttgac ctttggtact 5280
actctcttcc gatgatgatg tcgcacttat tctatgctgt ctcaatgtta gaggcatatc 5340
agtctccact gaagccaatc tatctgtgac ggcattcttta ttcacattat cttgtacaaa 5400
taatcctgtt aacaatgctt ttatatcctg taaagaatcc attttcaaaa tcatgtcaag 5460
gtcttctcga ggaaaaatca gtagaaatag ctgttccagt ctttctagcc ttgattccac 5520
ttctgtcaga tgtgccctag tcagcggaga ccttttgggt ttgggagagt agcgacactc 5580
ccagttgttc ttcagacact tggcgcaact cggtttttct ttggagcact tgagcttttt 5640
aagtcggcaa atatcgcatg cttgttcgat agaagacagt agcttcatgg tggcggaatcc 5700
cgataagcca gtaagcagtg ggttctctag ttagccagag agctctgctt atatagacct 5760
cccaccgtac acgcctaccg cccatttgcg tcaatggggc ggagttgtta cgacattttg 5820
gaaagtcctg ttgattttgg tgccaaaaca aactccatt gacgtcaatg ggggtggagac 5880
ttggaaatcc ccgtgagtc aaccgctatc cacgcccatt gatgtactgc caaaaccgca 5940
tcaccatggt aatagcgatg actaatacgt agatgtactg ccaagtagga aagtcccata 6000
aggtcagtga ctgggcataa tgccaggcgg gccatttacc gtcattgacg tcaatagggg 6060
gcgtacttgg catatgatac acttgatgta ctgccaaagt ggagttttac cgtaaatact 6120
ccacccattg acgtcaatgg aaagtcccta ttggcgttac tatgggaaca tacgtcatta 6180
ttgacgtcaa tgggcggggg tcgttggggc gtcagccagg cggggccattt accgtaagtt 6240
atgtaacgcg gaactccata tatgggctat gaactaatga ccccgtaatt gattactatt 6300
aataactagt caataatcaa tgtcaacgcg tatactctggc ccgtacatcg cgaagcagcg 6360
caaaacgcct aaccctaccc gggctgcagg aattctaccg ggtaggggag gcgcttttcc 6420
caaggcagtc tggagcatgc gcttttagcag ccccgctggc acttggcgca tcacaagtgg 6480
cctctggcct cgcacacatt ccacatccac cggtagcgcc aaccggctcc cttctttggg 6540
ggccccttgc cgccaccttc tactcctccc ctagttagga agttcccccc cgccccgcag 6600
ctcgcgtcgt gcaggacgtg acaaatggaa gtagcacgtc tcactagtct cgtgcagatg 6660
gacaagcacc gctgagcaat ggaagcgggt aggcctttgg ggcagcggcc aatagcagct 6720
tggctccttc gctttctggg ctccagaggc ggggaagggg ggggtccggg gcgggctcag 6780
gggcgggctc agggggcggg cgggcgcgaa ggtcctccgg tcttctcat ctccgggctc tcgacctgca 6900
caaaagcgca cgtctgccgc gctgttctcc gcgacgtctg tcgagaagtt tctgatcgaa aagttcgaca 6960
tgaaaaagcc tgaactcacc ggcagctctg ctctcggagg gcgaagaatc tcgtgctttc agcttcgatg 7020
gcgtctccga cctgatgcag ctctcggagg ggcgaagaatc cgatggtttc taaaagatc 7080
taggagggcg tggatatgtc ctgcgggtaa atagctgcgc cgatggtttc taaaagatc 7140
gttatgttta tcggcacttt gcacgcggcg cgctcccgat tccggaagtg cttgacattg 7200
gggaattcag cgagagcctg acctattgca tctcccgcg tgcacagggt gtcacgttgc 7260
aagacctgcc tgaaccgaa ctgcccgctg ttctgcagcc ggtcgcggag gccatggatg 7320
cgatcgctgc ggcgatctt agccagacga gcgggttcgg cccattcgga ccgcaaggaa 7354
tcggtcaata cactacatgg cgtgatttca tatg

```

&lt;210&gt; 50

&lt;211&gt; 10514

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence: Synthetic recombinant  
DNA vector sequence

&lt;400&gt; 50

```

ctaaattgta agcgttaata ttttggttaa attcgcgtta aatttttggg aaatcagctc 60
atthttttaac caataggccg aaatcggcaa aatcccttat aaatcaaaaag aatagaccga 120
gatagggttg agtggtgttc cagtttgtaa caagagtcca ctattaaaga acgtggactc 180
caacgtcaaa gggcgaaaaa ccgtctatca gggcgatggc ccactacgtg aaccatcacc 240
ctaatacaagt tttttggggg cgaggtgccc taaagcacta aatcggaacc ctaaagggag 300
cccccgattt agagcttgac ggggaaagcc ggcgaacgtg gcgagaaaag aagggaagaa 360
agcgaaagga gcggggcgcta gggcgctggc aagtgtagcg gtcacgctgc gcgtaaccac 420
cacacccgcc gcgcttaatg cgccgctaca gggcgcgctc cattcgccat tcaggctgcg 480
caactgttgg gaagggcgat cggtgcgggc ctcttcgcta ttacgccagc tggcgaaagg 540

```

gggatgtgct	gcaaggcgat	taagttgggt	aacgccaggg	ttttcccagt	cacgacgttg	600
taaaacgacg	gccagtgaac	gcgcgtaata	cgactcacta	tagggcggaat	tggagctcca	660
ccgcggtgcg	gccgggccat	gcaggccaga	catgataaga	tacattgatg	agtttggaca	720
aaccacaact	agaatgcagt	gaaaaaaatg	ctttatttgt	gaaatttgtg	atgctattgc	780
tttatttcta	accattataa	gctgcaataa	acaagttgtc	gagttacttg	tacagctcgt	840
ccatgccgag	agtgatcccc	gcggcggtca	cgaactccag	caggaccatg	tgatcgcgct	900
tctcgttggg	gtctttgctc	agggcggtac	ggtagctcag	gtagtgggtg	tcgggcagca	960
gcacggggcc	gtcgcgcatg	gggggtgtct	gctggtagtg	gtcggcgagc	tgacgctcgc	1020
cgtcctcgat	gttgtggcgg	atcttgaagt	tcaccttgat	gccgttcttc	tgcttgcctg	1080
ccatgatata	gacgttgtgg	ctgttgtagt	tgtactccag	cttgtgcccc	aggatgttgc	1140
cgtcctcctt	gaagtcatg	cccttcagct	cgatgcggtt	caccagggtg	tcgcccctga	1200
acttcacctc	ggcgcggttc	ttgtagtgtc	cgtcgtcctt	gaagaagatg	gtgcgctcct	1260
ggacgtagcc	ttcgggcatg	gcggacttga	agaagtcgtg	ctgcttcatg	tggctcgggtg	1320
agcgggagaa	gcactgcagg	ccgtagccga	aggtggtcac	gaggggtggc	cagggcacgg	1380
gcagcttgcc	ggtggtgcag	atgaacttca	gggtcagctt	gccgtagggt	gcacgccttc	1440
cgcctcgcgc	ggacacgctg	aacttgtggc	gctttacgtc	gccgtccagc	tcgaccagga	1500
tgggcaccac	cccgtgaac	agctcctcgc	ccttgcctac	catggtggcg	ttagcttgat	1560
ttgacagtgg	ctgggggttg	cgccgcgggg	ttttatagga	agccacagcg	gccactcgag	1620
ccataaaaagg	caactttagg	aacggcgggg	ggtgattgga	ttcgagtcgt	ttattcaccg	1680
gccttgccgc	acagtgcagc	atttttttac	ccccctctcc	ctccttttgc	gggggaaaaa	1740
aaaaaaaaaa	aaaaaaaaag	gagaagagaa	aaaaagcgag	cgagagagaa	agcgagattg	1800
aggaagagga	tgaagagttt	ggcgatgggt	gctggttccg	tagggccaga	tggacaagaa	1860
tagccccgcg	ccttgccgac	agtatcccat	tcagtgactc	agatcagtc	agcggccgcg	1920
tctaacctag	gagcccacca	cagacactat	actataacca	cactgaattt	tctctcctat	1980
ctcaaataat	cctattcttg	ttcatgcctc	agtgccttat	aacaagcaga	tttctcggcc	2040
agacatgggt	gcttgccgct	gtaatcccag	cactttggcc	aagtcgggca	gctcacttga	2100
ggtcaggagt	ttgaattcag	cctgtccaat	atggtgaaac	cccattctta	ctaaaaatac	2160
aaaaattagc	tgggcatggt	ggtgtgtgcc	tgtagtccca	gctactcagg	agactgaggc	2220
aggagaattg	cttgaactcg	ggaagcagag	ggtgcagtga	gccaagatca	tgccattgca	2280
ctccagcctg	ggcaacagag	cgagactccg	tctcaaacaa	aacaaaacaa	aaacaaaaaa	2340
acaagcagat	ctctcaactt	ggaatgctgc	tgcttctcca	tgccaactgg	aaaacttctt	2400
tttcttttag	actcaactca	aatatgatcc	tccttggaat	ctctccatct	ttcctagaca	2460
caattagcaa	ccagttcctc	tgtgttcaaa	cagcactctg	tgtgtatgtg	tatgtatgtg	2520
tatttatcat	tattattatc	acctttgatc	actctcctg	aaggtaatag	ccaggcactt	2580
agtatagagt	aggatatcaac	aattgtctat	tgaaggataa	aattgggtgaa	tgacaacttt	2640
gcaaagtttc	aatttttaaat	ctattgattt	gaaagaatta	ttatacatta	ttttaaatat	2700
cctttctagt	tctaatactt	tatgtctata	gtagttcgac	agttacttta	aatcagatat	2760
taaatgattt	tcattttaca	aagaacagac	ctgcctttct	caatgtcaca	agactattgt	2820
agagaatagc	agtgaataaa	gtacatgtga	aaatacttgg	caaagtaata	gtgaaaatag	2880
agctgatctc	acaggaaatg	tttgaaaatt	aaaacaccgt	atgtaaaata	ggtaatctaa	2940
tctgtgggca	catattttaa	ctctcaataa	atattaactt	tgattattta	aaaaatacct	3000
ggtagctttt	ttgtgaattc	tacaagaacc	tgtacatgat	tggttgccat	ttcctgcaaa	3060
atgagaaaat	tttcttctgc	actgaattct	tcttttaact	aaatttttaag	aaaaaaacca	3120
atgattaaaa	agtagataag	aaaatgataa	tataaaacca	ttagactaaa	gaaaaactag	3180
gttttagatga	taaaatttga	aaacccatct	tactatttct	cagatctttg	gagggcaatt	3240
accattatca	agatagggtg	gagacttcat	actcttgaat	tggttaaatg	tgctgagagt	3300
ttaaaagtct	aacttgaatt	ctcagctcta	atcatttcaa	agtaacatat	tcacatataa	3360
attatacctc	caaggcaata	aaaactccag	aaaatattaa	ctttaaaccc	acataattttc	3420
agatattata	catacaattt	tatttggtat	ttcctgaggc	atcctctgtt	tggttatatga	3480
atccataata	aaatgtagaa	gagtcgtgtg	atctgggggtg	agttcagttt	tctcctacac	3540
tggcaagaaa	ataaaaagggtg	ttagggaaga	tgtgtataaa	gccttctttc	tatatgtgca	3600
ttcatctgga	gggacctaat	agcagaaatc	ataaaaaaaa	aaaggcataa	atacagcagg	3660
aactatgagt	tctaagagag	tgtataattt	cttctcttgg	gtcctcaggc	aggccatatc	3720
acatttcact	gaattttcct	atctataaaa	tatgaataac	aaaatattcc	acttgtgtat	3780
ttcttaggaa	taaagtgtag	tctggatatg	aagactattg	agccatttca	tattgaagggt	3840
cactgaaagt	tcaagcttta	attataacta	atgaaagggt	ttcatatagt	atgactacta	3900
tgaatactat	attcatatac	tattctacat	gaattccacc	tggtgaagtc	atctgagagt	3960
attcatgtgg	aggaggggatc	aacaaactgt	ggcccatgga	ccaaatccat	tggattttat	4020

aaagttttat	tggaacacag	tcacactcat	ttgtttatgt	attgtctctg	ctactttttgc	4080
actaaaatgg	cagagttgag	tagctgttgt	agataccatg	tggccaaagc	ccttttaaaga	4140
aaaaaagttt	gtcaaccctt	aaaatagaga	gtcaacaatc	tggacaattc	agaccccaaa	4200
catctctttc	tccctctaca	gcatcttttg	aagagtgtga	ataattatgt	caactctttg	4260
ttttcttata	ccaaagtgtt	actgatgaag	cactatggca	aaataatgct	ttcatataaa	4320
atactttttc	ttaaataaat	acaaaaatac	gtattttgtaa	aactaaatgt	ggataaagtt	4380
ttctttgtca	ataatgtaat	aaatattgag	aaggggtgtg	tttggacaat	catagcttac	4440
tgggtttttca	gctaattggca	aaataaaaat	ccatcaaatg	tcctgtgcat	tatcacaatg	4500
tgtgtgtcta	aacaccttgg	atgctgttcc	cctattttgt	cctaagttgc	tcaatggatt	4560
tcttgagaaa	tttgtaaccc	cagtgggtag	gacttttctc	tattcatcca	catacgtgta	4620
ttcacccctg	tttactcatt	tataaggatt	aatccacagc	ctctattgag	accatgttct	4680
ttccagtttg	agactgtctt	ctcaaaaaga	gggaaaaaat	ggatcatcaat	actcagttgg	4740
aaaaaatcca	ttttatacac	caaatcattt	ctgctggcaa	acactaggta	gtaggactgg	4800
aatggagtc	gagagaagga	aaactttcac	tttttagttt	atgcattgct	actatttgaa	4860
tgtttacatc	aaagcatgta	ttaggaatcc	attttaatga	acaattggag	tagaaattca	4920
gctgtgctga	tggctttaat	ttgaaagcct	cctggccagc	tcactatgta	cctagaaact	4980
cagttccctg	cagtcttggc	agacaccatt	gcatattatt	accctgcatg	actttgttgt	5040
cgaggctcact	tgtcgcaagt	cacgaccttc	actgtcttct	ctagacggcg	atacagtcaa	5100
ctgtctttga	cctttgttac	tactctcttc	cgatgatgat	gtcgcactta	ttctatgctg	5160
tctcaatggt	agaggcata	cagtctccac	tgaagccaat	ctatctgtga	cggcatcttt	5220
attcacatta	tcttgtacaa	ataatcctgt	taacaatgct	tttatatcct	gtaaagaatc	5280
cattttccaa	atcatgtcaa	ggtcttctcg	aggaaaaatc	agtagaaata	gctgttccag	5340
tctttctagc	cttgattcca	cttctgtcag	atgtgacctt	gtcagcggag	accttttggg	5400
tttgggagag	tagcgacact	cccagttgtt	cttcagacac	ttggcgact	tcggtttttc	5460
tttggagcac	ttgagctttt	taagtcggca	aatatcgcat	gcttggtcga	tagaagacag	5520
tagcttcatg	gtggcggatc	ccgataagcc	agtaagcagt	gggttctcta	gttagccaga	5580
gagctctgct	tatatagacc	tcccaccgta	cacgectacc	gcccatttgc	gtcaatgggg	5640
cggagttggt	acgacatttt	ggaaagtccc	gttgattttg	gtgccaaaac	aaactcccat	5700
tgacgtcaat	gggtgggaga	cttgggaaatc	cccgtagatc	aaaccgctat	ccacgcccac	5760
tgatgtactg	ccaaaaccgc	atcaccatgg	taatagcgat	gactaatacg	tagatgtact	5820
gccaaagtag	aaagtcccat	aaggatcatg	actgggcata	atgccaggcg	ggccattttac	5880
cgatcatgac	gtcaataggg	ggcgactctg	gcatatgata	cacttgatgt	actgccaaat	5940
gggcagttta	ccgtaaatag	tccacccatt	gacgtcaatg	gaaagtccct	attggcggtta	6000
ctatgggaac	atacgtcatt	attgacgtca	atgggcgggg	gtcgttgggg	ggtcagccag	6060
cggggccatt	taccgttaagt	tatgtaacgc	ggaactccat	atatgggcta	tgaactaatg	6120
accccgtaat	tgattactat	taataactag	tcaataatca	atgtcaacgc	gtatatctgg	6180
cccgatcatc	gcgaagcagc	gcaaaacgcc	taacctacc	cgggctgcag	gaattctacc	6240
gggtaggagg	ggcgcttttc	ccaaggcagt	ctggagcatg	cgcttttagca	gccccgctgg	6300
cacttggcgc	atcacaaagt	gcctctggcc	tcgcacacat	tccacatcca	ccggtagcgc	6360
caaccggctc	ccttcttttg	tggccccttc	gcgccacctt	ctactctctc	cctagtccag	6420
aagttccccc	ccgcccgcga	gctcgcgctg	tgcaggacgt	gacaaatgga	agtagcacgt	6480
ctcactagtc	tcgtgcagat	ggacaagcac	cgctgagcaa	tggaaagcgg	taggcctttg	6540
gggcagcggc	caatagcagc	ttggctcctt	cgctttctgg	gctcagaggc	tgggaagggg	6600
tgggtccggg	ggcgggctca	ggggcgggct	caggggcggg	gcgggcgcga	aggtcctccg	6660
gacccggcat	tctgcacgct	tcaaaaagcg	acgtctgccg	cgctgttctc	ctcttctctca	6720
tctccggggc	ttcgacctgc	atgaaaaagc	ctgaactcac	cgcgacgtct	gtcgagaagt	6780
ttctgatcga	aaagtctcgac	agcgtctccg	acctgatgca	gctctcgga	ggcgaagaat	6840
ctcgtgcttt	cagcttcgat	gtaggagggg	gtggatatgt	cctgcgggta	aatagctgcg	6900
ccgatgggtt	ctacaaagat	cgttatgttt	atcggcactt	tgcacgggcc	gcgctcccga	6960
ttccggaagt	gcttgacatt	ggggaattca	gcgagagcct	gacctattgc	atctcccgcc	7020
gtgcacaggg	tgtcacgttg	caagacctgc	ctgaaaccga	actgcccgtc	gttctgcagc	7080
cggtcgcgga	ggccatggat	gcgatcgctg	cggccgatct	tagccagacg	agcgggttcg	7140
gccatttcgg	accgaagga	atcgggtcaat	acactacatg	gcgtgatttc	atatgcgcga	7200
ttgctgatcc	ccatgtgtat	cactggcaca	ctgtgatgga	cgacaccgtc	agtgcgtccg	7260
tcgcgcaggg	tctcgatgag	ctgatgcttt	gggcccaggga	ctgccccgaa	gtccggcacc	7320
tcgtgcacgc	ggatttcggc	tccaacaatg	tcttgacgga	caatggccgc	ataacagcgg	7380
tcattgactg	gagcgaggcg	atgttcgggg	attcccaata	cgaggtcgcc	aacatcttct	7440
tctggaggcc	gtgggttggct	tgtatggagc	agcagacgcg	ctacttcgag	cggaggcatc	7500



cgagagcttgc	aggatcgccg	cggtccggg	cgtatatgct	ccgcattggt	cttgaccaac	7560
tctatcagag	cttggttgac	ggcaatttcg	atgatgcagc	ttgggcgag	ggcgcgatgcg	7620
acgcaatcgt	ccgatccgga	gccgggactg	tcgggcgtag	acaaatcgcc	cgcagaagcg	7680
cgccggtctg	gaccgatggc	tgtgtagaag	tactcgccga	tagtggaac	cgacgcccc	7740
gcactcgctc	gagggcaaag	gaatagagta	tgcagaaatt	gatgatctat	taaacaataa	7800
agatgtccac	taaaatggaa	gtttttcctg	tcatactttg	ttaagaaggg	tgagaacaga	7860
gtacctacat	tttgaatgga	aggattggag	ctacgggggt	gggggtgggg	tggtattaga	7920
taaatgcctg	ctctttactg	aaggctcttt	actattgctt	tatgataatg	tttcatagtt	7980
ggatatcata	atttaaaca	gcaaaaccaa	attaagggcc	agctcattcc	tcccactcat	8040
gatctataga	tctatagatc	tctcgtggga	tcattgtttt	tctcttgatt	cccactttgt	8100
ggttctaagt	actgtggttt	ccaaatgtgt	cagtttcata	gcctgaagaa	cgagatcagc	8160
agcctctgtt	ccacatacac	ttcattctca	gtattgtttt	gccaagtctt	aattccatca	8220
gatttttagtg	gacatcttta	ttgtttaata	gatcatcaat	ttctgcatac	taagcttatc	8280
gataccgtcg	acctcgaggg	ggggcccggt	acccagcttt	tgttcccttt	agtgagggtt	8340
aattgcgcgc	ttggcgtaat	catggtcata	gctgtttcct	gtgtgaaatt	gttatccgct	8400
cacaattcca	cacaacatac	gagccggaag	cataaagtgt	aaagcctggg	gtgcctaattg	8460
agtgaactaa	ctcacattaa	ttgcgttgcg	ctcactgccc	gctttccagt	cgggaaacct	8520
gtcgtgccag	ctgcattaat	gaatcgccca	acgcgcgggg	agaggcggtt	tgcgtattgg	8580
gcgctcttcc	gcttcctcgc	tcactgactc	gctgcgctcg	gtcgttcggc	tgcggcgagc	8640
ggatatcagct	cactcaaagg	cggtataacg	gttatccaca	gaatcagggg	ataacgcagg	8700
aaagaacatg	tgagcaaaa	gccagcaaaa	ggccaggaac	cgtaaaaagg	ccgcgttgct	8760
ggcggtttttc	cataggctcc	gccccctga	cgagcatcac	aaaaatcgac	gctcaagtca	8820
gaggtggcga	aaccgcacag	gactataaag	ataccaggcg	tttccccctg	gaagctccct	8880
cgtgcgctct	cctgttccga	ccctgccgct	taccggatac	ctgtccgctt	ttctcccttc	8940
gggaagcgtg	gcgctttctc	atagctcacg	ctgtagggtat	ctcagttcgg	tgtaggctcg	9000
tcgctccaag	ctgggctgtg	tgacgaacc	ccccgttcag	cccagaccgt	gcgccttatc	9060
cggttaactat	cgtcttgagt	ccaacccggt	aagacacgac	ttatcgccac	tggcagcagc	9120
cactggtaac	aggattagca	gagcgaggta	tgtaggcggt	gctacagagt	tcttgaagtg	9180
gtggcctaac	tacggctaca	ctagaaggac	agtatttggt	atctgcgctc	tgctgaagcc	9240
agttaccttc	ggaaaaagag	ttggtagctc	ttgatccggc	aaacaaacca	ccgctggtag	9300
cggtgggtttt	tttgtttgca	agcagcagat	tacgcgcaga	aaaaaaggat	ctcaagaaga	9360
tcctttgatc	ttttctacgg	ggtctgacgc	tcagtggaac	gaaaactcac	gttaagggat	9420
tttggtcatg	agattatcaa	aaaggatctt	cacctagatc	cttttaaatt	aaaaatgaag	9480
ttttaaatca	atctaaagta	tatatgagta	aacttggtct	gacagttacc	aatgcttaat	9540
cagtgaaggca	cctatctcag	cgatctgtct	atttcgttca	tccatagttg	cctgactccc	9600
cgtcgtgtag	ataactacga	tacgggaggg	cttaccatct	ggccccagtg	ctgcaatgat	9660
accgcgagac	ccacgctcac	cggtccaga	tttatcagca	ataaaccagc	cagccggaag	9720
ggccgagcgc	agaagtggtc	ctgcaacttt	atccgcctcc	atccagtcta	ttaattgttg	9780
ccgggaagct	agagtaagta	gttcgccagt	taatagtttg	cgcaacgttg	ttgccattgc	9840
tacaggcatc	gtgggtgcac	gctcgtcggt	tggtatggct	tcattcagct	ccggttccca	9900
acgatcaagg	cgagttacat	gatcccccat	gttgtgcaaa	aaagcggtta	gctccttcgg	9960
tcctccgatc	gttgctcagaa	gtaagtggc	cgcagtgtta	tcactcatgg	ttatggcagc	10020
actgcataat	tctcttactg	tcatgccatc	cgtaagatgc	ttttctgtga	ctgggtagta	10080
ctcaaccaag	tcattctgag	aatagtgtat	gcggcgaccg	agttgctctt	gcccggcgtc	10140
aatacgggat	aataccgcgc	cacatagcag	aactttaaaa	gtgctcatca	ttggaaaacg	10200
ttcttcgggg	cgaaaactct	caaggatctt	accgctgttg	agatccagtt	cgatgtaacc	10260
cactcgtgca	cccaactgat	cttcagcatc	ttttactttc	accagcgttt	ctgggtgagc	10320
aaaaacagga	aggcaaaatg	ccgcaaaaaa	gggaataagg	gcgacacgga	aatgttgaat	10380
actcatactc	ttcctttttc	aatattattg	aagcatttat	cagggttatt	gtctcatgag	10440
cggatacata	tttgaatgta	tttagaaaaa	taaacaaata	ggggttccgc	gcacatttcc	10500
ccgaaaagtg	ccac					10514

&lt;210&gt; 51

&lt;211&gt; 10247

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Synthetic recombinant DNA molecule

&lt;220&gt;

&lt;221&gt; modified\_base

&lt;222&gt; (2626)..(2626)

&lt;223&gt; unsequenced nucleotide

&lt;400&gt; 51

ctagaactag	tggatccgga	taagccagta	agcagtgggt	tctctagtta	gccagagagc	60
tctgcttata	tagacctccc	accgtacacg	cctaccgccc	atttgcgctca	atggggcgga	120
gttggttacga	catttttgaa	agtcccggtg	atthttggtgc	caaaacaaac	tcccattgac	180
gtcaatgggg	tggagacttg	gaaatccccg	tgagtcaaac	cgctatccac	gcccattgat	240
gtactgcaa	aaccgcatca	ccatggtaat	agcgatgact	aatacgtaga	tgtactgcca	300
agtaggaaa	tcccataaag	tcatgtactg	ggcataatgc	caggcgggcc	atttaccgtc	360
attgacgtca	atagggggcg	tacttggcat	atgatacact	tgatgtactg	ccaagtgggc	420
agtttaccgt	aaatactcca	cccattgacg	tcaatggaaa	gtccctattg	gcgttactat	480
gggaacatac	gtcattattg	acgtcaatgg	gcgggggtcg	ttgggcgggc	agccaggcgg	540
gccatttacc	gtaagtatat	taacgcggaa	ctccatatat	gggctatgaa	ctaattgaccc	600
cgtaattgat	tactattaat	aactagtcaa	taatcaatgt	caaccccggg	ctgcaggaat	660
tctaccgggt	aggggaggcg	cttttcccaa	ggcagtcagg	agcatgcgct	ttagcagccc	720
cgctggcact	tggcgcatca	caagtggcct	ctggcctcgc	acacattcca	catccaccgg	780
tagcgccaac	cggtccctt	ctttgggtggc	cccttcgcgc	caccttctac	tcctccccta	840
gtcaggaagt	tccccccgcg	cccgcagctc	gcgtcgtgca	ggacgtgaca	aatggaagta	900
gcacgtctca	ctagtctcgt	gcagatggac	aagcaccgct	gagcaatgga	agcgggtagg	960
cctttggggc	agcggccaat	agcagcttgg	ctccttcgct	ttctgggctc	agaggctggg	1020
aaggggtggg	tccgggggcg	ggctcagggg	cgggctcagg	ggcggggcg	gcgcgaaggt	1080
cctccggacc	cggcattctg	cacgcttcaa	aagcgcacgt	ctgccgcgct	gttctcctct	1140
tcctcatctc	cgggccttcg	acctgcatga	aaaagcctga	actcaccgcg	acgtctgtcg	1200
agaagtttct	gatcgaaaag	ttcgacagcg	tctccgacct	gatgcagctc	tcggagggcg	1260
aagaatctcg	tgctttcagc	ttcgatgtag	gagggcggtg	atatgtcctg	cgggtaaata	1320
gctgcgccga	tggtttctac	aaagatcggt	atgtttatcg	gcactttgca	tcggccgcgc	1380
tcccgatctc	ggaagtgtct	gacattgggg	aattcagcga	gagcctgacc	tattgcatct	1440
cccgcggtgc	acaggggtgc	acgttgcaag	acctgcctga	aaccgaactg	cccgtgtgtc	1500
tgcagccggt	cgcggaggcc	atggatgcga	tcgctgcggc	cgatcttagc	cagacgagcg	1560
ggttcggccc	attcggaccg	caaggaatcg	gtcaatacac	tacatggcgt	gatttcatat	1620
gcgcgattgc	tgatcccat	gtgtatcact	ggcaaactgt	gatggacgac	accgtcagtg	1680
cgcccgctcg	gcaggctctc	gatgagctga	tgttttgggc	cgaggactgc	cccgaagtcc	1740
ggcacctcgt	gcacgcggat	ttcggctcca	acaatgtcct	gacggacaat	ggccgcataa	1800
cagcggctcat	tactggagc	gaggcgatgt	tcggggattc	ccaatacgag	gtcgcgaaca	1860
tcttcttctg	gaggccgtgg	ttggcttgta	tggagcagca	gacgcgctac	ttcgagcgga	1920
ggcatccgga	gcttgcaagg	tcgccgcggc	tcggggcgta	tatgtccgc	attggtcttg	1980
accaactcta	tcagagcttg	gttgacggca	atthcgatga	tgacgcttgg	gcgcaggggtc	2040
gatgcgacgc	aatcgtccga	tccggagccg	ggactgtcgg	gcgtacacaa	atcgcccgcga	2100
gaagcgcggc	cgtctggacc	gatggctgtg	tagaagtact	cgccgatagt	ggaaaccgac	2160
gccccagcac	tcgtccgagg	gcaaagggaat	gcctgagaaa	ggaagtgagc	tgtaaaaggct	2220
gagctctctc	tctgacgtat	gtagcctctg	gttagcttcg	tcactcactg	ttcttgactc	2280
agcatggcaa	tctgatgaaa	tcccagctgt	aagtctgcat	aaattgatga	tctattaaac	2340
aataaagatg	tccactaaaa	tggaagttht	ttactgtcat	actttgttaa	gaagggtgag	2400
aacagagtac	ctacatthttg	aatggaagga	ttggagctac	gggggtgggg	gtgggggtgg	2460
gattagataa	atgcctgctc	tttactgaag	gctctttact	attgtcttat	gataatgttt	2520
catagttgga	tatcataatt	taaacaagca	aaaccaatt	aagggccagc	tcattcctcc	2580
actcatgata	tatagatcta	tagatctctc	gtgggatcat	tgthtntctg	atccactgga	2640
agcttatcga	taccgtcgac	ctcgaggggg	ggcccggtac	ccagctthttg	ttcccttttag	2700
tgagggttaa	ttgcgcgctt	ggcgtaatca	tggtcatagc	tgthtctctg	gtgaaattgt	2760
tatccgctca	caattccaca	caacatacga	gccggaagca	taaagtgtaa	agcctgggggt	2820
gcctaattgag	tgagctaact	cacattaatt	gcgttgcgct	cactgcccgc	tttccagctcg	2880

ggaaacctgt	cgtgccagct	gcattaatga	atcggccaac	gcgcggggag	aggcggtttg	2940
cgtattgggc	gctcttcgcg	ttcctcgctc	actgactcgc	tgcgctcggg	cgttcggctg	3000
cggcgagcgg	tatcagctca	ctcaaaggcg	gtaatacggg	tatccacaga	atcaggggat	3060
aacgcaggaa	agaacatgtg	agcaaaaggc	cagcaaaagg	ccaggaaccg	taaaaaggcc	3120
gcgttgctgg	cgtttttcca	taggctccgc	ccccctgacg	agcatcacaa	aaatcgacgc	3180
tcaagtcaga	ggtggcgaaa	cccgcagga	ctataaagat	accaggcggt	ttccccctgga	3240
agctccctcg	tgcgctctcc	tgttccgacc	ctgcogetta	ccggatacct	gtccgccttt	3300
ctcccttcgg	gaagcgtggc	gctttctcat	agctcacgct	gtaggtatct	cagttcgggtg	3360
taggtcgttc	gctccaagct	gggctgtgtg	cacgaacccc	ccgttcagcc	cgaccgctgc	3420
gccttatccg	gtaactatcg	tcttgagtcg	aaccgggtaa	gacacgactt	atcgccactg	3480
gcagcagcca	ctggtaacag	gattagcaga	gcgagggtatg	taggcgggtgc	tacagagttc	3540
ttgaagtggg	ggcctaacta	cggctacact	agaaggacag	tatttggtat	ctgcgctctg	3600
ctgaagccag	ttaccttcgg	aaaaagagtt	ggtagctcct	gatccggcaa	acaaaccacc	3660
gctggtagcg	gtgggtttttt	tgtttgcaag	cagcagatta	cgcgcagaaa	aaaaggatct	3720
caagaagatc	ctttgatctt	ttctacgggg	tctgacgctc	agtggaaacga	aaactcacgt	3780
taaggggattt	tggcatgag	attatcaaaa	aggatcttca	cctagatcct	tttaaattaa	3840
aaatgaagtt	ttaaatcaat	ctaaagtata	tatgagtaaa	cttgggtctga	cagttacca	3900
tgcttaatca	gtgaggcacc	tatctcagcg	atctgtctat	ttcgttcatc	catagttgcc	3960
tgactccccg	tcgtgtagat	aactacgata	cgggaggggt	taccatctgg	ccccagtgct	4020
gcaatgatac	cgcgagaccc	acgctcaccg	gctccagatt	tatcagcaat	aaaccagcca	4080
gccggaaggg	ccgagcgcag	aagtggctct	gcaactttat	ccgcctccat	ccagtctatt	4140
aattgttgcc	gggaagctag	agtaagtagt	tgccaggtta	atagtttgcg	caacgttggt	4200
gaccttgcta	caggcatcgt	ggtgtcacgc	tcgtcgtttg	gtatggcttc	attcagctcc	4260
ggttcccaac	gatcaaggcg	agttacatga	ttccccatgt	tgtgcaaaaa	agcgggttagc	4320
tccttcgggtc	ctccgatcgt	tgtcagaagt	aagttggccg	cagtggttatc	actcatgggt	4380
atggcagcac	tgcataattc	tcttactgtc	atgccatccg	taagatgctt	ttctgtgact	4440
ggtgagtact	caaccaagtc	attctgagaa	tagtgtatgc	ggcgaccgag	ttgctcttgc	4500
ccggcgctcaa	tacgggataa	taccgcgcga	atagcagaa	ctttaaaagt	gctcatcatt	4560
ggaaaacggt	cttcggggcg	aaaactctca	aggatcttac	cgtgtgtgag	atccagttcg	4620
atgtaaccca	ctcgtgcacc	caactgatct	tcagcatctt	ttactttcac	cagcgtttct	4680
gggtgagcaa	aaacaggaag	gcaaaatgcc	gcaaaaaagg	gaataagggc	gacacggaaa	4740
tgttgaatac	tcatactctt	cctttttcaa	tattattgaa	gcatttatca	gggttattgt	4800
ctcatgagcg	gatacatatt	tgaatgtatt	tagaaaaata	aacaaatagg	ggttccgcgc	4860
acatttcccc	gaaaagtgcc	acctaatttg	taagcgttaa	tattttgtta	aaattcgctg	4920
taaattttttg	ttaaatcagc	tcatttttta	accaataggc	cgaatccggc	aaaatccctt	4980
ataaatcaaa	agaatagacc	gagatagggt	tgagtgttgt	ttcagtttgg	aacaagagtc	5040
cactattaaa	gaacgtggac	tccaacgtca	aaggcgaaaa	aaccgtctat	cagggcgatg	5100
gccactacg	tgaaccatca	ccctaataca	gttttttggg	gtcgaggtgc	cgtaaagcac	5160
taaactcgaa	ccctaaaggg	agcccccgat	ttagagcttg	acggggaaaag	ccggcgaaacg	5220
tggcgagaaa	ggaagggaag	aaagcgaaa	gagcggggcg	tagggcgctg	gcaagtgtag	5280
cggtcacgct	gcgcgtaacc	accacaccgc	ccgcgcttaa	tgcgcgcgta	cagggcgctg	5340
cccattcgcc	attcaggctg	cgcaactgtt	gggaaggggcg	atcgggtgcg	gcctcttcgc	5400
tattacgcca	gctggcgaaa	gggggatgtg	ctgcaaggcg	attaagttgg	gtaacgccag	5460
ggttttccca	gtcacgacgt	tgtaaaacga	cggccagtga	gcgcgcgtaa	tacgactcac	5520
tatagggcga	attggagctc	caccgcgggtg	cggccggggcc	atgcaggcca	cgacatgata	5580
agatacatgt	atgagtttgg	acaaaccaca	actagaatgc	agtgaaaaaa	atgctttatt	5640
tgtgaaattt	gtgatgctat	tgcttttatt	gtaaccatta	taagctgcaa	taaacaaagt	5700
cttgtagcgc	tcgtccatgc	cgagagtgat	ccggcgggcg	gtcacgaact	ccagcaggac	5760
catgtgatcg	cgcttctcgt	tggggtcttt	gctcagggcg	gactggtagc	tcaggtagtg	5820
gttgctcggg	agcagcacgg	ggccgtcgcc	gatgggggtg	ttctgctggg	agtggtcggc	5880
gagctgcacg	ctgccgtcct	cgatgtttgtg	gcggatcttg	aagttcacct	tgatgccgtt	5940
cttctgcttg	tcggccatga	tatagacgtt	gtggctgttg	tagttgtact	ccagctttgtg	6000
ccccaggtg	ttgccgtcct	ccttgaagtc	gatgcccttc	agctcgatgc	ggttccaccg	6060
gggtgcgccc	tcgaacttca	cctcggcgcg	ggctcttgtag	ttgccgtcgt	ccttgaagaa	6120
gatggtgcgc	tcctggacgt	agccttcggg	catggcggac	ttgaagaagt	cgtgctgctt	6180
catgtggtcg	gggtagcggg	cgaagcactg	caggccgtag	ccgaagggtg	tcacgagggt	6240
gggccagggc	acgggcagct	tgccgggtgg	gcagatgaac	ttcaggggtca	gcttgccgta	6300
gggtggcatcg	ccctcgccct	cgcgggacac	gctgaacttg	tggccggtta	cgtcgccgctc	6360

cagctcgacc	aggatgggca	ccaccccggt	gaacagctcc	tgcgcccttgc	taccatgggt	6420
ggcggttagct	tgatttgaca	gtggctgggg	gttgccgcgc	cgggttttat	aggaagccac	6480
agcggccact	cgagccataa	aaggcaactt	taggaacggc	ggggggtgat	tggattcgag	6540
tcgtttatct	accggccttg	ccgcacagtg	cagcattttt	ttaccccttc	tcccctcctt	6600
ttgcggggga	aaaaaaaaaa	aaaaaaaaaa	aaaggagaag	agaaaaaaag	cgagcgagag	6660
agaaagcgag	attgaggaag	aggatgaaga	gtttggcgat	gggtgctggg	tccgtaggcc	6720
cagatggaca	agaatagccc	ccgcccttgc	ggacagtatc	ccattcagtg	actcagatca	6780
gatcaagcgg	ccgccagtgt	gatggatatc	tgcagaattc	cagcacactg	gcggccgtta	6840
ctagtggatc	ctcaatcagt	ccatcacctg	ggtacttttt	ctacccaatg	ctgaaagtga	6900
atatttgaga	aactatgttt	ctaactggaa	gaattctgaa	actatgttct	tctacaacat	6960
gtagacgatg	tgtcacaatg	aaaatgaaag	ttccttctga	agagctgtaa	ccctacatag	7020
cctaggatgc	catttatact	ttctctagcc	cagaggctct	gtatttctgc	tctcaggcta	7080
cattactaca	taaccatcaa	accaactgcc	cgaaattgct	gagcaattga	tcataacata	7140
cagacctgat	tcagatatat	aaattacatt	tttattttta	tcgtaacaaa	acatacaggt	7200
gaaaactttc	aaattcaggc	aactgacatg	ccagggtgct	acctatctgt	gaagaccaga	7260
gcactgtaag	acacggaata	gctggttgaa	aagtgaaaaa	tccagtgagt	taggttgttt	7320
gattatacag	aatcatagaa	ttttcaaaag	gaaaggacat	gggtgatcact	gatttcagtg	7380
actatccaac	aacattagca	agagccccta	gggttcccta	ccagttcctt	cagaagcctc	7440
cttcaaagct	gtgagtgcac	ggccaaaggg	gaagctacat	tcataacca	gactaatcag	7500
agaccagggg	gcagaaataa	ggttaaacacg	gaatggagta	catttgcttt	gggggacgtt	7560
catgttttct	tgcgcatact	gaaaacagct	gggtgctgat	tttgggtataa	cactcagagc	7620
ctcttttgat	aaaaaggcat	tctatatgaa	tgaaaaccct	gttgggttca	catttggtcaa	7680
agcaatggat	tttaaatagt	gactcaactt	cctagaaaaa	tttgtattta	gaaatgtttt	7740
cttcatacag	ctaacaggca	aaactaatca	aaaccttggt	gttagtctga	tgtttatttt	7800
ctgtgcagca	aacaaactcc	cataaaaacc	tctgagggtt	tcttagttat	gtttgcaatt	7860
ttttaagtgt	tgaactgcct	aagtctctga	tgatcatgaca	cttaattcat	atattttctt	7920
gcttttgcta	gaactgaaca	aacactccct	atcaagattg	accctctcta	gaaataaaaat	7980
aaattgcatc	atgaagtcaa	aatatgacct	atcaagttta	tgatagcttt	ctttgacatt	8040
tatgggagga	gaaaattata	ccatccatac	tacatagggt	gggggtttaa	taatatttgc	8100
ctataaacia	taagcaattg	cagattgatt	ttaagtaaac	gtatgcaatg	cctgtaagca	8160
acaggaaata	tattttacat	caatgtatgc	actgcctgta	agtaacagga	attatatattt	8220
acatgaatgt	gtgaagtgat	aattttttta	taggcaaaaa	ttagtgttaa	aagcggaaact	8280
ggcatgcttt	cctaggctgt	ctcagtgata	atgacctcaa	agtcaatgat	ttcttcatga	8340
agatgaaata	agtttattgc	aacttgattg	catcacctaa	aatacaacc	ttgacttggg	8400
agcttgccct	tatgttttat	gaggtctgaa	ataaaaagga	aactgatgga	actgaaccag	8460
caaattatct	tgcaatccat	gactacagtt	gtatacagca	gatgtagtca	attgttgagg	8520
aaagttggac	caaatttcca	tcctgttcag	tactgctgat	agctctattc	aaaatacatt	8580
cttgcccttta	gaggaaactga	cccattagta	atcatacaga	tagaattttg	tgaactatctt	8640
ttaaagggta	ggttattaat	caatgagtga	ctctcaaaata	tcttgatggg	ttccactgtg	8700
aaaaccaagt	atctgcatag	tagagcaaaa	atttaatttt	ctcagatgct	ttagagtaca	8760
tatagatgaa	gactgccttc	ttactagatc	agtttttgaga	gagatgaaaa	caaaagaaag	8820
aaaaaatagg	atagctgatg	aaatgaaact	tcaaatttta	tttgatcatg	ttggcatttt	8880
ttatttcctg	agtgcattga	aaataactga	ttaatactag	acacataata	tactatcaag	8940
ttacacctaa	aatttatctt	gtcagccagg	caaccaatca	gaaccaataa	gtgaccaaga	9000
attatttttt	caaacacggt	tgtcttccag	gaccttactg	agttagtctc	acagatcatg	9060
aaaacatggg	tcttttctca	ggtcttctac	atcatttccc	ttaagttgac	acccaaagtc	9120
ataagcttga	ctttagacac	agggtttcaaa	agtagaaagt	tttattctct	tatgtttcta	9180
tggttttgaa	gttttgacga	actataaaaa	taatttggtt	tacccttttc	acaactcatg	9240
gacgttcctt	gaagtgaagg	aaacgaatac	tgcgtcttgc	ataataaaca	gcaaaaaaaa	9300
aaaaaaaaaa	aaaattaaat	gaatctaaat	tctcatatgc	catttaaagt	ttcagaccca	9360
gacatcttca	ggattttctt	acatttaaag	cagaatgaca	ctactgatac	acaaaagtaa	9420
atttttaattc	tgtttttaat	gcatgaggca	atataaaaaat	taatcatgat	gatcattata	9480
tttgttctaa	cactaaagtt	tttcccataa	tccccagat	ttttcaagtc	tcttttctcc	9540
tcatggcaga	tgcccaacca	cagagagaaa	ggttctataa	taaaacttta	aaaacactga	9600
cctaaaagag	aaatgtttta	tcgccatgca	atattctaaa	ttctggcatt	tggtggaaaa	9660
actgtacaca	tttttgcaa	tggtcttagt	gtccactggg	ctggcagcta	taatgagaat	9720
aataaagtat	atgggtgatga	ctgtttactt	accattatga	gacatcccca	ctgcgggtacc	9780
cggcgatata	gtcaactgtc	tttgaccttt	gttactactc	tcttccgatg	atgatgtcgc	9840

acttatttcta	tgctgtctca	atgttagagg	catatcagtc	tccactgaag	ccaatctatc	9900
tgtgacggca	tctttattca	cattatcttg	tacaaataat	cctgttaaca	atgcttttat	9960
atcctgtaaa	gaatccattt	tcaaaatcat	gtcaaggctc	tctcgaggaa	aaatcagtag	10020
aaatagctgt	tccagtcttt	ctagccttga	ttccacttct	gtcagatgtg	ccctagtcag	10080
cggagacctt	ttggttttgg	gagagtagcg	acactcccag	ttgttcttca	gacacttggc	10140
gcacttcggt	ttttctttgg	agcacttgag	ctttttaagt	cggcaaatat	cgcagtcttg	10200
ttcgatagaa	gacagtagct	tcatggtggc	aagcttaagt	ttaaacg		10247

&lt;210&gt; 52

&lt;211&gt; 10514

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Synthetic recombinant DNA vector sequence

&lt;400&gt; 52

ctaaattgta	agcgttaata	ttttgttaaa	attcgcgtta	aatttttgtt	aaatcagctc	60
attttttaac	caataggccg	aaatcggcaa	aatcccttat	aaatcaaaaag	aatagaccga	120
gatagggttg	agtgttggtc	cagtttgtaa	caagagtcca	ctattaaaga	acgtggactc	180
caacgtcaaa	ggcgaaaaa	ccgtctatca	ggcgatggc	ccactacgtg	aaccatcacc	240
ctaatacaagt	tttttggggg	cgaggtgccg	taaagcacta	aatcggaacc	ctaaagggag	300
cccccgattt	agagcttgac	ggggaaaagg	ggcgaaacgtg	gcgagaaaagg	aagggaagaa	360
agcgaaaagga	gcgggcgcta	gggcgctggc	aagtgtagcg	gtcacgctgc	gcgtaaccac	420
cacaccggcc	gcgcttaatg	cgccgctaca	gggcgcgtcc	cattcgccat	tcaggctgcg	480
caactgttgg	gaagggcgat	cggcgcgggc	ctcttcgcta	ttacgccagc	tggcgaaaagg	540
gggatgtgct	gcaaggcgat	taagtgggg	aacgccaggg	ttttcccagt	cacgacgttg	600
taaaacgacg	gccagtgcag	gcgcgtaata	cgactcacta	tagggcgaa	tggagctcca	660
ccgcggtgcg	gccgggccat	gcaggccaga	catgataaga	tacattgatg	agtttggaca	720
aaccacaact	agaatgcagt	gaaaaaaatg	ctttattttgt	gaaattttgtg	atgctattgc	780
tttattttgta	accattataa	gctgcaataa	acaagttgtc	gagttacttg	tacagctcgt	840
ccatgccgag	agtgatcccg	gcggcggtca	cgaactccag	caggaccatg	tgatcgcgct	900
tctcgttggg	gtctttgctc	agggcgggact	ggtagctcag	gtagtgggtg	tcgggcagca	960
gcacggggcc	gtcgccgatg	ggggtgttct	gctggtagtg	gtcggcgagc	tgcacgctgc	1020
cgctcctcgat	gttggtggcg	atcttgaagt	tcaccttgat	gccgttcttc	tgcttgtcgg	1080
ccatgatata	gacgttggtg	ctgttgtagt	tgtactccag	cttggtgccc	aggatgttgc	1140
cgctcctcctt	gaagtgcagt	cccttcagct	cgatgcgggt	caccaggggtg	tcgcccctcga	1200
acttcacctc	ggcgcggtc	ttgtagttgc	cgctgcctc	gaagaagatg	gtgcgtcctc	1260
ggacgtagcc	ttcgggcgatg	gcggacttga	agaagtcgtg	ctgcttcagt	tggtcggggg	1320
agcggggcgaa	gcactgcagg	ccgtagccga	aggtggtcac	gaggggtggg	caggggcacgg	1380
gcagcttgcc	ggtggtgcag	atgaacttca	gggtcagctt	gccgtaggtg	gcatcgccct	1440
cgccctcgcc	ggacacgctg	aacttggtgg	cgtttacgtc	gccgtccagc	tcgaccagga	1500
tgggcaccac	cccgtgtaac	agctcctcgc	ccttgctcac	catggtggcg	ttagcttgat	1560
ttgacagtgg	ctgggggttg	cgccgcgggg	ttttatagga	agccacagcg	gccactcgag	1620
ccataaaaagg	caacttttagg	aacggcgggg	ggtgattgga	ttcgagtcgt	ttattcaccg	1680
gccttgccgc	acagtgcagc	atttttttac	ccccctctcc	ctccttttgc	gggggaaaaa	1740
aaaaaaaaaa	aaaaaaaaaag	gagaagagaa	aaaaagcgag	cgagagagaa	agcgagattg	1800
aggaagagga	tgaagagttt	ggcgatgggt	gctggttccg	taggcccaga	tggacaagaa	1860
tagccccgcg	ccttgcgagc	agtatcccat	tcagtgcactc	agatcagtc	agcggccgcg	1920
cttaacctag	gagcccacca	cagacactat	actataacca	cactgaattt	tctctcctat	1980
tcataataat	cctattcttg	ttcatgcctc	agtgccttat	aacaagcaga	tttctcggcc	2040
agacatggtg	gcttgcgcct	gtaatcccag	cactttggcc	aagtcgggca	gctcacttga	2100
ggtcaggagt	ttgaattcag	cctgtccaat	atggtgaaac	cccatctcta	ctaaaaatac	2160
aaaaattagc	tgggcatggt	ggtgtgtgcc	tgtagtccca	gctactcagg	agactgaggc	2220
aggagaattg	cttgaactcg	ggaagcagag	ggtgcagtga	gccaagatca	tgccattgca	2280
ctccagcctg	ggcaacagag	cgagactccg	tctcaaacaa	aacaaaaaaa	aaacaaaaaa	2340

acaagcagat	ctctcaactt	ggaatgctgc	tgcttctcca	tgccaactgg	aaaacttctt	2400
tttcttttag	actcaactca	aatatgatcc	tccctgaagt	ctctccatct	ttcctagaca	2460
caattagcaa	ccagttcctc	tgtgttcaaa	cagcacttcg	tgtgtatgtg	tatgtatgtg	2520
tatttatcat	tattattatc	acctttgatc	actctccctg	aaggtaatag	ccaggcactt	2580
agtatagagt	aggtatcaac	aattgcttat	tgaaggaata	aatgggtgaa	tgacaacttt	2640
gcaaagtttc	aatttttaac	ctattgattt	gaaagaatta	ttatacatta	ttttaaatat	2700
cctttctagt	tctaatatct	tatgtctata	gtagttcgac	agttacttta	aatcagatat	2760
taaatgattt	tcatttttaca	aagaacagac	ctgcctttct	caatgtcaca	agactattgt	2820
agagaatagc	agtgaataaa	gtacatgtga	aaatacttgg	caaagtaata	gtgaaaatag	2880
agctgatctc	acaggaatgt	tttgaaaatt	aaaacaccgt	atgtaaaata	ggtaactctaa	2940
tctgtgggca	catattttaa	ctctcaataa	atattaactt	tgattattta	aaaaaacctt	3000
ggtagctttt	ttgtgaattc	tacaagaacc	tgtacatgat	tgggtgccat	ttccgtcaaa	3060
atgagaaaaa	tttcttctgc	actgaattct	tcttttaact	aaatttttaag	aaaaaaacca	3120
atgattaaaa	agtagataag	aaaatgataa	tataaaacca	ttagactaaa	gaaaaactag	3180
gttttagatga	taaaatttga	aaacccatct	tactatttct	cagatctttg	gagggcaatt	3240
accattatca	agatagggtg	gagacttcat	actcttgaat	tgtttaaatg	tgctgagagt	3300
ttaaaagtct	aacttgaatt	ctcagctcta	atcattttcaa	agtaacatat	tcacataata	3360
attatacctc	caaggcaata	aaaactccag	aaaatattaa	ctttaaaccc	acatattttc	3420
agatattata	catacaattt	tatttggtat	ttcctgaggc	atcctctggt	tgttatatga	3480
atccataata	aatgtagaaa	gagtctgttg	atctgggggtg	agttcagttt	tctcctacac	3540
tggcaagaaa	ataaaaggtg	ttagggaaga	tgtgtataaa	gccttctttc	tatatgtgca	3600
ttcatctgga	gggacctaat	agcagaaatc	ataaaaaaaa	aaaggcataa	atacagcagg	3660
aacatagagt	tctaagagag	tgtataattt	ctttctctgg	gtcctcaggc	aggccatctc	3720
acatttctact	gaattttcct	atctataaaa	tatgaataac	aaaatattcc	acttgtgtat	3780
ttcttaggaa	taaatggtag	tctggatatg	aagactattg	agcccattca	tattgaagggt	3840
cactgaaagt	tcaagcttta	attataacta	atgaaagggt	ttcatatagt	atgactacta	3900
tgaatactat	attcatatac	tattctacat	gaattccacc	tgggtgaagtc	atctgagagt	3960
attcatgttg	aggagggatc	aacaaaactgt	ggcccattga	ccaaatccat	tggattttat	4020
aaagttttat	tggaacacag	tcacactcat	ttgtttatgt	attgtctctg	ctacttttgc	4080
actaaaatgg	cagagttgag	tagctgttgt	agataccatg	tggccaaagc	cctttaaaga	4140
aaaaaagttt	gtcaaccctt	aaaatagaga	gtcaacaatc	tggacaattc	agaccccaaa	4200
catctctttc	tccctctaca	gcatcttttg	aagagtgtga	ataattatgt	caactccttg	4260
ttttcttata	ccaaagtgtt	actgatgaag	cactatggca	aaataatgct	ttcatataaa	4320
atactttttc	ttaataaaat	acaaaaatac	gtatttgtaa	aactaaatgt	ggataaagtt	4380
ttctttgtca	ataattgtaat	aaatattgag	aaggggtggg	tttggacaat	catagcttac	4440
tgggtttttca	gctaattggca	aaataaaaat	ccatcaaattg	tctgtgtgcat	tatcacaaatg	4500
tgtgtgtcta	aacaccttgg	atgctgttcc	cctattttgt	cctaagttgc	tcaatggatt	4560
tcttgagaaa	tttgtaaccc	cagtgggtag	gacttttctc	tattcatcca	catacgtgta	4620
ttcacccctgt	tttactcatt	tataaggatt	aatccacagc	ctctattgag	accatgttct	4680
tcccagtttg	agactgtcct	ctcaaaaaga	gggaaaaaat	gggtcatcaat	actcagttgg	4740
aaaaaatcca	ttttatacac	caaatcattt	ctgtgggcaa	acactaggta	gtaggactgg	4800
aatggagtca	gagagaagga	aaactttcac	tttttagttt	atgcattgct	actatttgaa	4860
tgtttacatc	aaagcatgta	ttaggaatcc	attttaatga	acaattggac	tagaaattca	4920
gctgtgcctg	tggctttaat	ttgaaagcct	cctggccagc	tcactatgta	cctagaaaact	4980
cagttcctgc	cagtcttggc	agacaccatt	gcatattatt	accctgcatg	actttgttgt	5040
cgaggtcact	tgtcgcaagt	cacgaccttc	actgtcttct	ctagacggcg	atacagtcaa	5100
ctgtctttga	cctttgttac	tactctcttc	cgatgatgat	gtcgcaactta	ttctatgctg	5160
tctcaatggt	agaggcatat	cagtctccac	tgaagccaat	ctatctgtga	cggcatcttt	5220
attcacatta	tcttgtacaa	ataatcctgt	taacaatgct	tttatatcct	gtaaagaatc	5280
cattttcaaa	atcatgtcaa	ggtcttctcg	aggaaaaatc	agtagaaata	gctgttccag	5340
tctttctagc	cttgattcca	cttctgtcag	atgtgcccta	gtcagcggag	accttttgggt	5400
tttgggagag	tagcgacact	cccagttggt	cttcagacac	ttggcgcact	tcggtttttc	5460
tttgagcac	ttgagctttt	taagtcggca	aatatcgcat	gcttgttcga	tagaagacag	5520
tagcttcatg	gtggcggatc	ccgataagcc	agtaagcagt	gggttctcta	gttagccaga	5580
gagctctgct	tatatagacc	tcccaccgta	cacgcctacc	gcccatttgc	gtcaatgggg	5640
cggagttggt	acgacatttt	ggaaagtccc	gttgattttg	gtgccaaaac	aaactcccat	5700
tgacgtcaat	ggggtggaga	cttggaatc	cccgtgagtc	aaaccgctat	ccacgcccac	5760
tgatgtactg	ccaaaaccgc	atcaccatgg	taatagcgat	gactaatacg	tagatgtact	5820

gccaaagtagg	aaagtcccat	aaggatcatgt	actggggcata	atgccaggcg	ggccattttac	5880
cgatcattgac	gtcaataggg	ggcgactttg	gcataatgata	cacttgatgt	actgccaagt	5940
gggcagtttta	ccgtaaaatac	tccacccatt	gacgtcaatg	gaaagtcctt	attggcggtta	6000
ctatggggaac	atacgtcatt	attgacgtca	atgggcgggg	gtcgttgggc	ggtcagccag	6060
gcggggccatt	taccgtaagt	tatgtaacgc	ggaactccat	atatgggcta	tgaactaatg	6120
accccgtaaat	tgattactat	taataactag	tcaataatca	atgtcaacgc	gtatatctgg	6180
cccgtacatc	gcgaagcagc	gcaaaacgcc	taaccctacc	cgggctgcag	gaattctacc	6240
gggtagggga	ggcgcttttc	ccaaggcagt	ctggagcatg	cgcttttagca	gccccgctgg	6300
cacttggcgc	atcacaagtg	gcctctggcg	tgcacacat	tccacatcca	ccggtagcgc	6360
caaccggctc	ccttcttttg	tggcccttc	gcgccacctt	ctactcctcc	cctagtcagg	6420
aagttccccc	ccgccccgca	gctcgcgctg	tgcaggacgt	gacaaatgga	agtagcacgt	6480
ctcactagtc	tcgtagcagat	ggacaagcac	cgctgagcaa	tggaaagcggg	taggcctttg	6540
gggcagcggc	caatagcagc	ttggctcctt	cgctttctgg	gctcagaggc	tgggaagggg	6600
tgggtccggg	ggcgggctca	ggggcgggct	caggggcggg	gcgggcgcga	aggtcctccg	6660
gaccgggcat	tctgcacgct	tcaaaaagcgc	acgtctgccg	cgctgttctc	ctcttctcca	6720
tctccgggct	tctgacctgc	atgaaaaagc	ctgaactcac	cgcgacgtct	gtcgaagaat	6780
ttctgactga	aaagtctgac	agcgtctccg	acctgatgca	gctctcggag	ggcgaagaat	6840
ctcgtgcttt	cagcttcgat	gtaggaggcg	gtggatatgt	cctgcgggta	aatagctgcg	6900
ccgatgggtt	ctacaaagat	cgttatgttt	atcggcactt	tgcacggcc	gcgctcccga	6960
ttccggaagt	gcttgacatt	ggggaattca	gcgagagcct	gacctattgc	atctcccgcc	7020
gtgcacaggg	tgtcacgttg	caagacctgc	ctgaaaccga	actgcccgt	gttctgcagc	7080
cggctcgcgga	ggccatggat	gcgatcgctg	cgcccgatct	tagccagacg	agcgggttcg	7140
gccatttcgg	accgcaagga	atcgggtcaat	acactacatg	gcgtgatttc	atatgcgcga	7200
ttgctgatcc	ccatgtgtat	cactggcaaa	ctgtgatgga	cgacaccgtc	agtgcgtccg	7260
tgcgcagggc	tctcgatgag	ctgatgcttt	gggcccaggga	ctgccccgaa	gtccggcacc	7320
tcgtagcagc	ggatttcggc	tccaacaatg	tcttgacgga	caatggccgc	ataacagcgg	7380
tcattgactg	gagcgaggcg	atgttcgggg	attcccaata	cgaggctcgcc	aacatcttct	7440
tctggaggcc	tgtgtgggct	tgtatggagc	agcagacgcg	ctacttcgag	cggaggcatc	7500
cggagcttgc	aggatcgccg	cggctccggg	cgtatattgt	ccgcatttgt	cctgaccaac	7560
tctatcagag	cttggttgac	ggcaatttcg	atgatgcagc	ttgggcccag	ggtcgatgcg	7620
acgcaatcgt	ccgatccgga	gccgggactg	tccggcgctac	acaaatcgcc	cgcagaagcg	7680
cggccgtctg	gaccgatggc	tgtgtagaag	tactcgccga	tagtggaaac	cgacgccccca	7740
gcactcgtcc	gagggc aaaag	gaatagagta	tgcagaaaatt	gatgatctat	taaaacaataa	7800
agatgtccac	taaaatggaa	gtttttcctg	tcatactttg	ttaagaaggg	tgagaacaga	7860
gtacctacat	tttgaatgga	aggattggag	ctacgggggt	gggggtgggg	tgggattaga	7920
taaatgcctg	ctctttactg	aaggctcttt	actattgctt	tatgataatg	tttcatagtt	7980
ggatatcata	atttaaacaa	gcaaaaccaa	attaagggcc	agctcattcc	tcccactcat	8040
gatctataga	tctatagatc	tctcgtggga	tcattgtttt	tctcttgatt	cccactttgt	8100
ggttctaagt	actgtggttt	ccaaatgtgt	cagtttcata	gcctgaagaa	cgagatcagc	8160
agcctctgtt	ccacatacac	ttcattctca	gtattgtttt	gccaagtctt	aattccatca	8220
gatttttagtg	gacatcttta	ttgtttaata	gatcatcaat	ttctgcatac	taagcttatc	8280
gataccgtcg	acctcgaggg	ggggcccggt	accagctttt	tgttcccttt	agtgagggtt	8340
aattgcgcgc	ttggcgtaat	catggtcata	gctgtttcct	gtgtgaaatt	gttatccgct	8400
cacaattcca	cacaacatac	gagccggaag	cataaagtgt	aaagcctggg	gtgcctaattg	8460
agtgagctaa	ctcacattaa	ttgcgttgcg	ctcactgccc	gctttccagt	cgggaaacct	8520
gtcgtgccag	ctgcattaat	gaatcgccca	acgcgcgggg	agaggcgggt	tgcgtattgg	8580
gcgctcttc	gcttctcgc	tcactgactc	gctgcgctcg	gtcgttcggc	tgcggcgagc	8640
ggatatcagct	cactcaaaagg	cggtaataacg	gttatccaca	gaatcagggg	ataacgcagg	8700
aaagaacatg	tgagcaaaaag	gccagcaaaa	ggccagggaac	cgtaaaaagg	ccgcgttgct	8760
ggcgtttttc	cataggctcc	gccccctga	cgagcatcac	aaaaatcgac	gctcaagtca	8820
gaggtggcga	aaccgcagacg	gactataaaag	ataccaggcg	tttccccctg	gaagctccct	8880
cgtgcgctct	cctgttccga	ccctgcgcgt	taccggatac	ctgtccgcct	ttctcccttc	8940
gggaagcgtg	gcgctttctc	atagctcagc	ctgtaggtat	ctcagttcgg	tgtaggtcgt	9000
tcggtccaag	ctgggctgtg	tgcacgaacc	ccccgttcag	cccagaccgt	gcgccttatc	9060
cggtaactat	cgtcttgagt	ccaaccgggt	aagacacgac	ttatcgccac	tggcagcagc	9120
cactggtaac	aggattagca	gagcgaggta	tgtaggcggt	gctacagagt	tcttgaagtg	9180
gtggcctaac	tacgggtaca	ctagaaggac	agtatttggt	atctgcgctc	tgctgaagcc	9240
agttaccttc	ggaaaaagag	ttggtagctc	ttgatccggc	aaacaaacca	ccgctggtag	9300

cggtgggtttt	tttgtttgca	agcagcagat	tacgcgcaga	aaaaaaggat	ctcaagaaga	9360
tcctttgatc	ttttctacgg	ggtctgacgc	tcagtggaaac	gaaaactcac	gttaagggat	9420
tttgggtcatg	agattatcaa	aaaggatctt	cacctagatc	cttttaaatt	aaaaatgaag	9480
ttttaaatca	atctaaagta	tatatgagta	aacttggtct	gacagttacc	aatgcttaat	9540
cagtgaggca	cctatctcag	cgatctgtct	atttcgttca	tccatagttg	cctgactccc	9600
cgtcgtgtag	ataactacga	tacgggaggg	cttaccatct	ggccccagtg	ctgcaatgat	9660
accgcgagac	ccacgctcac	cggctccaga	tttatcagca	ataaaccagc	cagccggaag	9720
ggccgagcgc	agaagtggtc	ctgcaacttt	atccgcctcc	atccagtcta	ttaattgttg	9780
ccgggaagct	agagtaagta	gttcgccagt	taatagtttg	cgcaacgttg	ttgccattgc	9840
tacaggcatc	gtggtgtcac	gctcgtcgtt	tggatggct	tcattcagct	ccggttccca	9900
acgatcaagg	cgagttacac	gatcccccat	gttggtgcaaa	aaagcgggta	gctccttcgg	9960
tcctccgatc	gttgtcagaa	gtaagtggc	cgcagtgtta	tactcatgg	ttatggcagc	10020
actgcataat	tctcttactg	tcatgccatc	cgtaagatgc	ttttctgtga	ctgggtgagta	10080
ctcaaccaag	tcattctgag	aatagtgtat	gcggcgaccg	agttgctctt	gcccggcgctc	10140
aatacgggat	aataccgcgc	cacatagcag	aactttaaaa	gtgctcatca	ttggaaaacg	10200
ttcttcgggg	cgaaaactct	caaggatctt	accgctgttg	agatccagtt	cgatgtaacc	10260
cactcgtgca	cccaactgat	cttcagcatc	ttttactttc	accagcgttt	ctgggtgagc	10320
aaaaacagga	aggcaaaatg	ccgcaaaaaa	gggaataagg	gcgacacgga	aatgttgaat	10380
actcatactc	ttcctttttc	aatattattg	aagcatttat	caggggttatt	gtctcatgag	10440
cggatacata	tttgaatgta	tttagaaaaa	taaacaaata	gggggtccgc	gcacatttcc	10500
ccgaaaagtg	ccac					10514

<210> 53  
 <211> 168  
 <212> DNA  
 <213> Homo sapiens

<400> 53						
cggaggacag	tactccgctc	ggaggacagt	actccgctcg	gaggacagta	ctccgctcga	60
ggacagtact	ccgctcggag	gacagtactc	cgatccgctc	actctagtcc	gctcggagga	120
cagtactccg	ctcggaggac	agtactccgc	tcgaggagca	gtactccg		168

<210> 54  
 <211> 22  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: Synthetic primer

<400> 54						
gagaaccac	tgcttactgg	ct				22

<210> 55  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: Synthetic primer

<400> 55						
caagatcgcc	ctcgctttg					20



<210> 56  
<211> 33  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic primer

<400> 56  
atatggtacc gcagtgggga tgtctcataa tgg 33

<210> 57  
<211> 30  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic primer

<400> 57  
aattggatcc tcaatcagtc catcacctgg 30

<210> 58  
<211> 30  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic primer

<400> 58  
cgtctagaga gtgtgtcctg tcagaagaac 30

<210> 59  
<211> 31  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic primer

<400> 59  
atgcggccgc actcctgacc tcaggtgatc c 31

<210> 60  
<211> 29  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic primer

<400> 60  
cgtctagaga agacagtga ggtcgtgac 29

<210> 61  
 <211> 29  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: Synthetic primer

<400> 61  
 gcgcggccgc gtctaacctt ggagcccac 29

<210> 62  
 <211> 28  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: Synthetic primer

<400> 62  
 gctctagagg acgaatgccca agatctga 28

<210> 63  
 <211> 29  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: Synthetic primer

<400> 63  
 caagcggccg ccagtgtgat ggatatctg 29

<210> 64  
 <211> 5218  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: Synthetic recombinant  
 DNA sequence

<400> 64  
 gccggagcgg tcgagttctg gaccgaccgg ctccgggttct cccgggactt cgtggaggac 60  
 gacttcgccg gtgtggtccg ggacgacgtg accctgttca tcagcgcggt ccaggaccag 120  
 gtggtgccgg acaacaccct ggcctgggtg tgggtgcgcg gcctggacga gctgtacgcc 180  
 gagtggctcg aggtcgtgtc cacgaacttc cgggacgcct cggggccggc catgaccgag 240  
 atcggcgagc agccgtgggg gcgggagttc gccttgcgcg acccgccggc caactgcgtg 300  
 cacttcgtgg ccgaggagca ggactgacac gtgctacgag atttcgattc caccgccgcc 360  
 ttctatgaaa ggttgggctt cggaatcgtt ttccgggacg cgggctggat gatcctccag 420  
 cgcggggatc tcatgctgga gttcttcgcc caccccaact tgtttattgc agcttataat 480  
 ggttacaaat aaagcaatag catcacaat ttcacaaata aagcattttt ttcactgcat 540  
 tctagttgtg gtttgtccaa actcatcaat gtatcttatt atgtctgtat accgtcgacc 600  
 tctagctaga gcttggcgta atcatggtca tagctgtttc ctgtgtgaaa ttgttatccg 660

ctcacaaatc	cacacaacat	acgagccgga	agcataaagt	gtaaagcctg	gggtgcctaa	720
tgagtgaact	aactcacatt	aattgcgttg	cgctcactgc	ccgctttcca	gtcgggaaac	780
ctgtcgtgcc	agctgcatta	atgaatcggc	caacgcgcgg	ggagaggcgg	tttgcgtatt	840
gggcgctctt	ccgcttcctc	gctcactgac	tcgctgcgct	cggtcgttcg	gctgcggcga	900
gcggtatcag	ctcactcaaa	ggcggtaata	cggttatcca	cagaatcagg	ggataacgca	960
ggaaagaaca	tgtgagcaaa	aggccagcaa	aaggccagga	accgtaaaaa	ggccgcgttg	1020
ctggcggttt	tccataggct	ccgccccct	gacgagcatc	acaaaaatcg	acgctcaagt	1080
cagaggtggc	gaaacccgac	aggactataa	agataccagg	cgttttcccc	tggagactcc	1140
ctcgtgcgct	ctcctgttcc	gaccctgccc	cttaccggat	acctgtccgc	ctttctccct	1200
tcgggaagcg	tggcgctttc	tcaatgctca	cgctgtaggt	atctcagttc	gggtgtaggtc	1260
gttcgctcca	agctgggctg	tgtgcacgaa	cccccgttc	agcccgaccg	ctgcgcctta	1320
tccggtaact	atcgtcttga	gtccaacccg	gtaagacacg	acttatcgcc	actggcagca	1380
gccactggta	acaggattag	cagagcgagg	tatgtagggc	gtgctacaga	gttcttgaag	1440
tggtggccta	actacggcta	cactagaagg	acagtatttg	gtatctgcgc	tctgctgaag	1500
ccagttacct	tcggaaaaag	agttggtagc	tcttgatccg	gcaaaacaaac	caccgctggt	1560
agtcctgttt	tttttgtttc	caagcagcag	attacgcgca	gaaaaaaaagg	atctcaagaa	1620
gatacctttg	tcttttctac	ggggtctgac	gctcagtgga	acgaaaactc	acgttaaggg	1680
attttgggtc	tgagattatc	aaaaaggatc	ttcacctaga	tcctttttaa	ttaaaaatga	1740
agttttaaat	caatctaaag	tatatatgag	taaacttggt	ctgacagtta	ccaatgctta	1800
atcagtgaag	cacctatctc	agcgatctgt	ctatttcggt	catccatagt	tgctgactc	1860
cccgctcgtg	agataactac	gatacgggag	ggcttaccat	ctggccccag	tgctgcaatg	1920
ataccgcgag	accacgcgtc	accggctcca	gatttatcag	caataaacca	gccaccgga	1980
agggccgagc	gcagaagtgg	tcctgcaact	ttatccgcct	ccatccagtc	tattaattgt	2040
tgccgggaag	ctagagtaag	tagttcgcca	gttaatatgt	tgcgcaacgt	tggtgccatt	2100
gctacaggca	tcgtggtgtc	acgctcgtcg	tttggtatgg	cttcattcag	ctccggttcc	2160
caacgatcaa	ggcgagttac	atgatcccc	atggtgtgca	aaaaagcggg	tagctccttc	2220
ggtcctccga	tcgttgtcag	aagtaagttg	gccgcagtg	tatcactcat	ggttatggca	2280
gactcgcata	attctcttac	tgtcatgcca	tcgctaagat	gcttttctgt	gacttcggag	2340
tactcaacca	agtcattctg	agaatagtgt	atgcggcgac	cgagttgctc	ttgcccggcg	2400
tcaatacggg	ataataccgc	gccacatagc	agaactttaa	aagtgtcat	catttgaaaa	2460
cgttcttcgg	ggcgaaaact	ctcaaggatc	ttaccgctgt	tgagatccag	ttcgatgtaa	2520
cccactcgtg	cacccaactg	atcttcagca	tcttttactt	tcaccagcgt	ttctgggtga	2580
gcaaaaacag	gaaggcaaaa	tgccgcaaaa	aagggaataa	gggcgacacg	gaaatgttga	2640
atactcatac	tcttcctttt	tcaatattat	tgaagcattt	atcagggtta	ttgtctcatg	2700
agcggtatac	tatttgaatg	tatttagaaa	aataaacaaa	taggggttcc	gcgcacattt	2760
ccccgaaaag	tgccacctga	cgtcgacgga	tcgggagatc	cgatgttttc	tgagttactt	2820
ttgtatcccc	acccccctc	gaggatcgat	ccgctcggag	gacagtactc	cgctcggagg	2880
acagtactcc	gctcggagga	cagtactccg	ctcggaggaca	gtactccgct	cggaggacag	2940
tactccgatc	cgtcgactct	agtccgctcg	gaggacagta	ctccgctcgg	aggacagtac	3000
tcgctcggga	ggacagtact	ccgctcggag	acagtactcg	attcgaggac	attttgacac	3060
cccataata	ttttccaga	attaacagta	ttaattgcat	ctcttggtca	agagttccct	3120
atcactctct	ttaatcacta	ctcacagtaa	cctcaactcc	tgcccaagct	tggtaccacc	3180
atggacccag	aaacgctggt	gaaagtaaaa	gatgctgaag	atcagttggg	tgacagagt	3240
ggttacatcg	aactggatct	caacagcggg	aagatccttg	agagttttcg	ccccgaagaa	3300
cgttttccaa	tgatgagcac	ttttaaagtt	ctgctatgtg	gcgcgggtatt	atcccgtatt	3360
gacgccgggc	aagagcaact	cggtcgccgc	atacactatt	ctcagaatga	cttggttgag	3420
tactcaccag	tcacagaaaa	gcatcttacg	gatggcatga	cagtaagaga	attatgcagt	3480
gctgccataa	ccatgagtga	taacactgcg	gccaaacttac	ttctgacaa	gatcggagga	3540
ccgaaggagc	taaccgcttt	tttgacaca	atgggggatc	atgtaactcg	ccttgatcgt	3600
tggaaccggg	agctgaatga	agccatacca	aacgacgagc	gtgacaccac	gatgcctgta	3660
gcaatggcaa	caacgttgcg	caaactatta	actggcgaa	tacttactct	agcttcccgg	3720
caacaattaa	tagactggat	ggaggcggat	aaagttgcag	gaccacttct	gcgctcggcc	3780
cttcggctg	cctggtttat	tgctgataaa	tctggagccg	gtgagcgtgg	gtctcgcggt	3840
atcattgcag	cactggggcc	agatggtaag	ccctcccgtg	tcgtagttat	ctacacgacg	3900
gggagtcagg	caactatgga	tgaacgaaat	agacagatcg	ctgagatagg	tgctcactg	3960
attaagcatt	ggtaatctag	agagggcccc	tttaaaccgc	ctgatcagcc	tcgactgtgc	4020
cttctagttg	ccagccatct	gttgtttgcc	cctcccccg	gccttccttg	accctggaag	4080
gtgccactcc	cactgtcctt	tcctaataaa	atgaggaaat	tgcacgcgat	tgtctgagta	4140

```

gggtgtcattc tattctgggg ggtgggggtgg ggcaggacag caaggggggag gattgggaag 4200
acaatagcag gcatgctggg gatgcggtgg gctctatggc ttctgaggcg gaaagaacca 4260
gctgggggtc taggggggtat cccacgcgc cctgtagcgg cgcattaagc gcggcgggtg 4320
tggtgggttac gcgcagcgtg accgctacac ttgccagcgc cctagcgcgc gtccttttcg 4380
ctttcttccc ttcttttctc gccacgttcg ccggttttcc cgtcaagct ctaaatacggg 4440
gcatcccttt agggttccga tttagtgtt tacggcacct cgaccccaaa aaacttgatt 4500
aggggtgatgg ttcacgtagt gggccatcgc cctgatagac ggttttttcgc ctttgacgt 4560
tggaagtcac gttctttaat agtggactct tgttccaaac tggaacaaca ctcaacccta 4620
tctcgggtcta ttcttttgat ttataaggga ttttggggat ttcggcctat tggttaaaaa 4680
atgagctgat ttaacaaaaa tttaacgcga attaatctg tggaatgtgt gtcagttagg 4740
gtgtggaaaag tccccaggct cccaggcag gcagaagtat gcaaagcatg catctcaatt 4800
agtcagcaac caggtgtgga aagtccccag gctccccagc aggcagaagt atgcaaagca 4860
tgcattctcaa ttagtcagca accatagtcc cgcccctaac tccgcccata ccgcccctaa 4920
ctccgcccag ttccgcccct tctccgcccc atgggtgact aatttttttt atttatgcag 4980
agggcggaggc cgctctgcc tctgagctat tccagaagta gtgaggaggc ttttttggag 5040
gcctaggctt ttgcaaaaag ctcccgggag cttgtatata cattttcgga tctgatcagc 5100
acgtgttgac aattaatcat cggcatagta tatcggcata gtataatacg acaaggtgag 5160
gaactaaacc atggccaagt tgaccagtgc cgttccgggtg ctcaccgcgc gcgacgtc 5218

```

<210> 65

<211> 28

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic primer

<400> 65

ggggtaccaa agaaggtagg ctgagggg

28

<210> 66

<211> 32

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic primer

<400> 66

cgggatccgt acaagacagt tagctagttg gc

32

<210> 67

<211> 34

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic primer

<400> 67

ggggtacat tcagcaggcc actacaggac tctc

34

<210> 68  
 <211> 32  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: Synthetic  
 primer

<400> 68  
 gatcgcgcc gctgtgctcg acattggtgg cc 32

<210> 69  
 <211> 31  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: Synthetic  
 primer

<400> 69  
 ggggtacctt tgtggtgctt aaaaaatgag c 31

<210> 70  
 <211> 33  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: Synthetic  
 primer

<400> 70  
 gatcgcgcc gcctcatgaa caatgaaatc tcc 33

<210> 71  
 <211> 2379  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: Synthetic  
 Nurrl targeting sequence

<400> 71  
 cgtacaagac agttagctag ttggcaaac caaggagaat tctgtgacaa gggaaactca 60  
 ggacaaatcc tgctaggcag ttctggaggg gaagcgccct gcgcctgcag tactgacctg 120  
 tgaccatagc cagggcagca atgcaggaga aggcagaaat gtcgatgttc atattctgca 180  
 agttggagga gaattcaaca atggaatcaa tccattcccc aaagccacga acgcattgca 240  
 acctgtgcaa gaccacccca ttgcaaaaga tgagtttacc ctccactggg ttggacctgc 300  
 aattaatacc aaagagagag aggggagaaa aagagagaga gaaaagctaa acaactaatt 360  
 acttgaagag caataaatga gggatttaag agtcacctaa ttactgaaga gttaattaaa 420  
 atgtagatcc agtggacctt gaaagggttt aatttcataa caatcaagaa cgccccctat 480  
 cccacctcca ttccattcat ccaaggcttc tgggctcgct tcttctcgtc tttttttggt 540

```

ttctttcttt ttcttcttt ctccgacttc catttcttat tctgtctttt tctctacccc 600
acctctgtgt ttcccttctt ccttttcttt tcttttcttg atttctctca cagcctccct 660
ggattgtctc ctccctccct tattacctgt atgctaatac aaggacaaac agttctaaga 720
aagctgattc aaaaagcagg tcttggtcgg ctttgggcag gtctgcgaag ccagggatct 780
tctctgcccc gccccggatg atctccatgg agccagtcag gagatcatag aattgctgga 840
tatgtctgggt gtcattctca ctcatattgat agtcagggtt cgcctggaac tggaaattca 900
ttttaaaaag cacttaataa ggttctctaa aatatataac ccgtgaaatt gctaaccctc 960
tttctaataa gggagccagg tttttataac aattaaacct ctctctgacc agtaaggaaa 1020
ttagatgttc cggggcaatt aattcaatta gggatgggtg tacgaggctg ctgctttaa 1080
tatgtaaatt gtcatttcca tacagactaa atacagtgc atccagccct gggaaacgtt 1140
cactcttctc tccacaaaac aattgacatg ttagtattca tgctagtccc agagtgggcc 1200
tcgagctgag acgccccttt gcaaacttta ataaaattgc agtccctacc aggtttgctc 1260
aacatacttc ctccaattaa agcccgcata tttttcacag ctggagtaaa aggatctgac 1320
cacttgatcc cccacccccc tggcatcttt acacctctc ctcccttcc ctgaccccca 1380
gacagccccc acatccttcc ccagcacagt gccaggaagg atgaggatta aagcatctat 1440
ctctatggag gactgggaat aaagtagatt ggggaagggg agagtttgca cagcccctgg 1500
atttcatggg agaggcactg gtgtatatgg aagaggaggg gcgatggggg agtggggtag 1560
tggaatttca cacagcttca aacctgtgg cgtgtccctc tctccttcc tccatctgt 1620
cctcatgtct atgccccag ctacatggat ttgtctgaag atctgaatct tgaggccctg 1680
gccagagctg cgaggcatat acagccttgc ttgcctttct ttacccctc tgaatctgag 1740
agttaatgac ggatgtgggg aggggtcctg cccatctgtc ggtttgtcca catgatattc 1800
ccccgcagc cttcttacc tggaaatgtc caggctgggtc atagccgggt tggagtgcac 1860
atgggccctg acgagggcac tgatcagact caccgggggc gaagggggag agggctcctg 1920
tggtctcttc ggtttcgagg gcaaacgacc tctccggcct tttaaactgt ctgtgcgaac 1980
cactgcaaag gaagagccct gttagcgcg cttttccgag cccaggccca gctgctgcct 2040
cggtccttcc ccggggaagg ccgcagccgc ggggcaccag gctgagcggc tgagggcccc 2100
agtgtctgta aagccttcac tgactagaag cattaataaa tgcggggtta ttttatgtct 2160
tcttccaaat gggtcgtata gttaaaggag agaaggccct ggcggctttc tctagggaag 2220
gccgggcaag caggcagctg cagggtcctg gaggccatac tgagggggag tcgagatcc 2280
ccagcaccgg gaagtgaac gtgatgctgg agtatgagca gtggtttcct aaaggcgcaa 2340
actggagggt cggcagctcc cctcagccta ccttctttg 2379

```

<210> 72

<211> 8426

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic  
pCDGal4-DBD-Nurrl sequence

<400> 72

```

tctctggcta actagagaac cactgctta ctggcttata gaaattaata cgactcacta 60
tagggagacc caagctgggt agcgtttaaa cttaagcttg ccaccatgaa gctactgtct 120
tctatcgaac aagcatgcga tatgtgccga cttaaaaagc tcaagtgtc caaagaaaa 180
ccgaagtgcg ccaagtgtct gaagaacaac tgggagtgtc gctactctcc caaaacaaa 240
aggtctccgc tgactagggt acatctgaca gaagtggat caaggctaga aagactggaa 300
cagctatttc tactgatttt tcctcgagaa gacctgaca tgattttgaa aatggattct 360
ttacaggata taaaagcatt gttaacagga ttatttgtac aagataatgt gaataaagat 420
gccgtcacag atagattggc ttcagtgagg actgatattg ctctaacatt gagacagcat 480
agaataagtg cgacatcatc atcggaagag agtagtaaca aaggtcaaag acagttgact 540
gtatcccggt gtaccaaaga aggtaggctg aggggagctg ccgacctcc agtttgcgcc 600
tttaggaac cactgctcat actccagcat cactgtccac tcccgggtgc tggggatctc 660
cgactcccc tcagtatggc ctccaggacc ctgcagctgc ctgcttgccc ggcttccct 720
agagaaagcc gccaggccct tctctccttt aactatacga cccatttgga ggaagacata 780
aaataacccc gcatttttta atgcttctag tcagtgaagg ctttacaagc actggggccc 840
tcagccgctc agcctgggtg cccgcggctg cggccttccc cggggaggga ccgaggcagc 900

```

agctgggcct	gggctcggaa	aagcggcgct	aacaggggctc	ttcctttgca	gtggttcgca	960
cagacagttt	aaaaggcccg	agaggtcggt	tgccttcgaa	accgaagagc	ccacaggagc	1020
cctctcccc	ttcgcccccg	gtgagtctga	tcagtgcctt	cgtcagggcc	catgtcgact	1080
ccaaccgcgc	tatgaccagc	ctggactatt	ccagggtaag	aagctggcgg	gggggatatc	1140
atgtggacaa	accgacagat	gggcaggacc	cctccccaca	tccgtcatta	actctcagat	1200
tcaacggggg	taaagaaagg	caagcaaggc	tgtatatgcc	tcgcagctct	ggccagggcc	1260
tcaagattca	gatcttcaga	caaattccatg	tagctggggg	catagacatg	aggacaggat	1320
ggaggaagga	ggagagggac	acgccacagg	gtttgaagct	gtgtgaattc	ccactacccc	1380
actaccccat	cgccccctct	cttccatata	caccagtgcc	tctaccatga	aatccagggg	1440
ctgtgcaaac	tctccccctt	cccaatctac	tttatcccca	gtcctccata	gagatagatg	1500
ctttaatcct	catccttcct	ggcactgtgc	tggggaagga	tgtgggggct	gtctgggggt	1560
caggaagggg	aaggagaggg	tgtaaagatg	ccagtggggg	gggggatcaa	gtggtcagat	1620
ccttttactc	cagctgtgaa	aaatatgcgg	gctttaattg	gaggaagtat	gttgagcaaa	1680
cctggtaggg	actgcaattt	tattaagatt	tgcaaaaggg	cgtctcagct	cgaggcccac	1740
tctgggacta	gcatgaatac	taacatgtca	attgttttgt	ggagataaga	gtgaacgttt	1800
cccagggctg	gatggcactg	tatttagtct	gtatggaaat	gacaattttac	atattttaag	1860
cagcgacctc	gtagcaccat	ccctaattga	attaattgcc	ccggaacatc	taatttcctt	1920
actggtcaga	gagaggttta	attgttataa	aaacctggct	cccctattag	aaacgggggt	1980
agcaattttca	cgggttatat	attttagaga	acctcattaa	gtgcttttta	aaatgaaatt	2040
ccagttccag	gcgaaccctg	actatcaaat	gagtggagat	gacaccagc	atatccagca	2100
attctatgat	ctcctgactg	gctccatgga	gatcatccgg	ggctggggcag	agaagatccc	2160
tggcttcgca	gacctgccca	aagccgacca	agacctgctt	tttgaatcag	ctttcttaga	2220
actgtttgtc	cttcgattag	catacaggta	ataagggagg	gaggagacaa	tccagggagg	2280
ctgtgagaga	aatcaagaaa	ggaaaagaaa	gggaggaagg	gaaaccagag	ggtggggtag	2340
agaaaaagac	agaataggaa	atggaagtcg	gagaaaggaa	gaaaaagaaa	gaaaacaaaa	2400
aaagacgaga	agaagcgagc	ccagaagcct	tggatgaatg	gaatggaggt	gggatagggg	2460
gcgttcttga	ttgttatgaa	attaaaccct	ttcaaggctc	actggatcta	catttttaatt	2520
aaactcttcag	taattaggtg	actcttaaat	ccctcattta	ttgctcttca	agtaattagt	2580
tgtttagctt	ttctctctct	ctttttctcc	cctctctctc	tttgggtatta	attgcaggctc	2640
caaccctagt	gagggtaaac	tcattctttt	caatgggggtg	gtcttgcaca	ggttgcaatg	2700
cgttcgtggc	tttggggaat	ggattgattc	cattgttgaa	ttctcctcca	acttgcaaaa	2760
tatgaacatc	gacatttctg	ccttctcctg	cattgctgcc	ctggctatgg	tcacagggtca	2820
gtactgcagg	cgcagggcgc	ttccccctcca	gaactgccta	gcaggatttg	tcctgagttt	2880
cccttgcac	agaattctcc	ttggttttgc	caactagcta	actgtcttgt	acggatccac	2940
tagtaacggc	gcgcagtggt	ctggaattct	cgagatatcc	atcacactgg	cggcgcgtcg	3000
agtctagagg	gcccgtttta	acccgctgat	cagcctcgac	tgtgccttct	agttgccagc	3060
catctgttgt	ttgccccctc	cccgtgcctt	ccttgacctt	ggaagggtgc	actcccactg	3120
tcctttccta	ataaaatgag	gaaattgcat	cgcatgtgtc	gagtaggtgt	cattctattc	3180
tgggggggtg	ggtggggcag	gacagcaagg	gggaggattg	ggaagacaat	agcaggcatg	3240
ctgggggatg	ggtgggctct	atggcttctg	aggcggaag	aaccagctgg	ggctctaggg	3300
ggtatcccca	cgcgccctgt	agcggcgcat	taagcgcggc	gggtgtgggt	gttacgcgca	3360
gcgtgaccgc	tacacttgcc	agcgccctag	cgccgcgtcc	tttcgctttc	ttccccctct	3420
ttctcgccac	gttcgcccgc	tttccccgtc	aagctctaaa	tcgggggctc	cctttagggt	3480
tcggatttag	tgctttacgg	cacctcgacc	ccaaaaaact	tgattagggt	gatggttcac	3540
gtagtgggct	atcgccctga	tagacgggtt	ttcgcccttt	gacgttggag	tcacagttct	3600
ttaatagtgg	actcttggtc	caaactggaa	caacactcaa	ccctatctcg	gtctattctt	3660
ttgatttata	agggattttg	ccgatttcgg	cctatttggtt	aaaaaatgag	ctgatttaac	3720
aaaaatttaa	cgcgaattaa	ttctgtggaa	tgtgtgtcag	ttagggtgtg	gaaagtcccc	3780
aggctcccca	gcaggcagaa	gtatgcaaag	catgcattct	aattagttag	caaccagggtg	3840
tggaaagtcc	ccaggctccc	cagcaggcag	aagtatgcaa	agcatgcatt	tcaatttagtc	3900
agcaaccata	gtcccgcccc	taactccgcc	catcccgccc	ctaactccgc	ccagttccgc	3960
ccattctccg	ccccatggct	gactaatttt	ttttattttt	gcagaggccg	aggccgcctc	4020
tcctcttgag	ctattccaga	agtagtgagg	aggctttttt	ggaggcctag	gcttttgcaa	4080
aaagctcccg	ggagcttgta	tatccatttt	cggatctgat	cagcacgtga	tgaaaaagcc	4140
tgaactcacc	gcgacgtctg	tcgagaagtt	tctgatcgaa	aagttcgaca	gcgtctccga	4200
cctgatgcag	ctctcggagg	gcgaagaatc	tcgtgctttc	agcttcgatg	taggagggcg	4260
tggatatgtc	ctgcgggtaa	atagctgcgc	cgatgggttc	tacaaagatc	gttatgttta	4320
tcggcacttt	gcatcgcccg	cgctcccgat	tccggaagtg	cttgacattg	gggaattcag	4380

cgagagcctg	acctattgca	tctcccgcg	tgcacaggg	gtcacgttgc	aagacctgcc	4440
tgaaccgaa	ctgcccgcg	ttctgcagcc	ggctgcggag	gccatggatg	cgatcgctgc	4500
ggccgatctt	agccagacga	gcgggttcgg	cccattcgga	ccgcaaggaa	tcgggtcaata	4560
cactacatgg	cgtgatttca	tatgcgcgat	tgctgatccc	catgtgtatc	actggcaaac	4620
tgtgatggac	gacaccgtca	gtgcgtccgt	cgcgagggct	ctcgatgagc	tgatgctttg	4680
ggccgaggac	tgccccgaag	tccggcacct	cgtgcacgcg	gatttcggct	ccaacaatgt	4740
cctgacggac	aatggccgca	taacagcgg	cattgactgg	agcgaggcga	tgttcgggga	4800
ttcccaatac	gaggtcgcca	acatcttctt	ctggaggccg	tgggttgctt	gtatggagca	4860
gcagacgcgc	tacttcgagc	ggaggcatcc	ggagcttgca	ggatcgccgc	ggctccgggc	4920
gtatatgctc	cgcattgggtc	ttgaccaact	ctatcagagc	ttgggtgacg	gcaatttcga	4980
tgatgcagct	tgggcgcagg	gtcgatgcga	cgcaatcgtc	cgatccggag	ccgggactgt	5040
cgggcgtaca	caaatcgccc	gcagaagcgc	ggccgtctgg	accgatggct	gtgtagaagt	5100
actcgccgat	agtggaaacc	gacgccccag	cactcgctcc	agggcaaagg	aatagcacgt	5160
gctacgagat	ttcgattcca	ccgccgcctt	ctatgaaagg	ttgggcttcg	gaatcgtttt	5220
ccgggacgcc	ggctggatga	tcctccagcg	cggggatctc	atgctggagt	tcttcgcccc	5280
ccccaaactg	tttattgcag	cttataatgg	ttacaaataa	agcaatagca	tcacaaattt	5340
cacaaataaa	gcattttttt	cactgcattc	tagttgtggg	ttgtccaaac	tcatcaatgt	5400
atcttatcat	gtctgtatac	cgtcgacctc	tagctagagc	ttggcgtaat	catggtcata	5460
gctgtttcct	gtgtgaaatt	gttatccgct	cacaattcca	cacaacatac	gagccggaag	5520
cataaagtgt	aaagcctggg	gtgcctaatt	agtgaagctaa	ctcacattaa	ttgcgttgcg	5580
ctcactgccc	gctttccagt	cgggaaacct	gtcgtgccag	ctgcattaat	gaatcggcca	5640
acgcgcgggg	agaggcgggt	tgcgtattgg	gcgctcttcc	gcttccctcg	tcactgactc	5700
gctgcgctcg	gtcgttcggc	tgcggcgagc	gttatcagct	cactcaaagg	cggtaatcag	5760
gttatccaca	gaatcagggg	ataacgcagg	aaagaacatg	tgagcaaaaag	gccagcaaaa	5820
ggccaggaac	cgtaaaaagg	ccgcgttgct	ggcgtttttc	cataggctcc	gccccctga	5880
cgagcatcac	aaaaatcgac	gctcaagtca	gagggtggcg	aacccgacag	gactataaag	5940
ataccaggcg	tttccccctg	gaagctccct	cgtgcgctct	cctgttccga	ccctgccgct	6000
taccggatac	ctgctccgct	ttctcccttc	gggaagcgtg	gcgctttctc	atagctcacg	6060
ctgtagggtat	ctcagttcgg	tgtaggctcg	tcgctccaag	ctgggctgtg	tgcacgaacc	6120
ccccgttcag	cccgaccgct	gcgccttata	cggtaaactat	cgtcttgagt	ccaacccggg	6180
aagacacgac	ttatcgccac	tggcagcagc	cactggtaac	aggattagca	gagcgaggta	6240
tgtaggcggg	gctacagagt	tcttgaagtg	gtggcctaac	tacggctaca	ctagaagaac	6300
agtatttggg	atctgcgctc	tgtgaagcc	agttaccttc	ggaaaaagag	ttggtagctc	6360
ttgatccggc	aaacaaacca	ccgctggtag	cggttttttt	gtttgcaagc	agcagattac	6420
gcgcagaaaa	aaaggatctc	aagaagatcc	tttgatcttt	tctacggggg	ctgacgctca	6480
gtggaacgaa	aactcacgtt	aagggtattt	ggatcatgaga	ttatcaaaaa	ggatcttcac	6540
ctagatcctt	ttaaattaaa	aatgaagttt	taaatcaatc	taaagtatat	atgagtaaac	6600
ttgggtctgac	agttaccaat	gcttaatcag	tgaggcacct	atctcagcga	tctgtctatt	6660
tcgttcaccc	atagttgcct	gactccccgt	cgtgtagata	actacgatac	gggaggggct	6720
accatctggc	cccagtgcgt	caatgatacc	gcgagaccca	cgctcaccgg	ctccagattt	6780
atcagcaata	aaccagccag	ccggaaggcg	cgagcgcaga	agtggctctg	caactttatc	6840
cgctccatc	cagtctatta	attgttgccg	ggaagctaga	gtaagtagtt	cgccagttaa	6900
tagtttgccg	aacgttgttg	ccattgctac	aggcatcggt	gtgtcacgct	cgtcgtttgg	6960
tatggcttca	ttcagctccg	gttcccaacg	atcaaggcga	gttacatgat	cccccatggt	7020
gtgcaaaaaa	gcgggttagct	ccttcgggtcc	tccgatcggt	gtcagaagta	agttggccgc	7080
agtgttatca	ctcatgggtta	tggcagcact	gcataattct	cttactgtca	tgccatccgt	7140
aagatgcctt	tctgtgactg	gtgagtactc	aaccaagtca	ttctgagaat	agtgtatgcg	7200
gcgaccgagt	tgtctttgcc	cggcgtcaat	acgggataat	accgcgccac	atagcagaac	7260
tttaaaagtg	ctcatcattg	gaaaacgttc	ttcggggcga	aaactctcaa	ggatcttacc	7320
gctgttgaga	tccagttcga	tgtaacccac	tctgtcaccc	aactgatctt	cagcatcttt	7380
tactttcacc	agcgtttctg	ggtgagcaaa	aacaggaagg	caaaatgccg	caaaaaaggg	7440
aataagggcg	acacggaaat	ggtgaatact	catactcttc	ctttttcaat	attattgaag	7500
catttatcag	ggttattgtc	tcatgagcgg	atacatattt	gaatgtattt	agaaaaataa	7560
acaaataggg	gttccgcgca	catttccccg	aaaagtgcc	cctgacgtcg	acggatcggg	7620
agatctcccg	atccccctatg	gtgcactctc	agtacaatct	gctctgatgc	cgcatagtta	7680
agccagtatc	tgtccctgc	ttgtgtgttg	gaggctcgct	agtagtgccg	gagcaaaatt	7740
taagctacaa	caaggcaagg	cttgaccgac	aattgcatga	agaatctgct	taggggttagg	7800
cgttttgccg	tgtctcgcga	tgtacggggc	agatatacgc	gttgacattg	attattgact	7860



```

agttattaat agtaatcaat tacggggtca ttagttcata gcccatatat ggagttccgc 7920
gttacataac ttacggtaaa tggcccgcct ggctgaccgc ccaacgaccc ccgcccattg 7980
acgtcaataa tgacgtatgt tcccatagta acgccaatag ggactttcca ttgacgtcaa 8040
tgggtggagt atttacggta aactgccac ttggcagtac atcaagtgt tcatatgcca 8100
agtacgcccc ctattgacgt caatgacggg aaatggcccc cctggcatta tgcccagtac 8160
atgaccttat gggactttcc tacttggcag tacatctacg tattagtcac cgctattacc 8220
atgggtgatgc ggttttggca gtacatcaat gggcgaggat agcggtttga ctcacgggga 8280
tttccaagtc tccaccccat tgacgtcaat gggagtttgt tttggcacca aaatcaacgg 8340
gactttccaa aatgtcgtaa caactccgcc ccattgacgc aaatgggcgg taggcgtgta 8400
cgggtgggagg tctatataag cagagc                                8426

```

<210> 73

<211> 9701

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic  
pKI-Gal4-DBD-Nurrl sequence

<220>

<221> modified\_base

<222> (1518)..(1518)

<223> any nucleotide

<400> 73

```

tgcacgcttc aaaagcgcac gtctgccgcg ctgtttctct cttcctcatc tccgggcctt 60
cgacctgcat gaaaaagcct gaactcaccg cgacgtctgt cgagaagttt ctgatcgaaa 120
agttcgacag cgtctccgac ctgatgcagc tctcggaggg cgaagaatct cgtgctttca 180
gcttcgatgt aggagggcgt ggatatgtcc tgcgggtaaa tagctgcgcc gatggtttct 240
acaaagatcg ttatgtttat cggcactttg catcgccgcg gctcccgaat ccggaagtgc 300
ttgacattgg ggaattcagc gagagcctga cctattgcat ctcccgcctg gcacaggggtg 360
tcacgttgca agacctgcct gaaaccgaac tgcccgcgtg tctgcagccg gtcgcggagg 420
ccatggatgc gatcgctgcg gccgatctta gccagacgag cgggttcggc ccattcggag 480
cgcaaggaat cggccaatac actacatggc gtgatttcat atgcgcgatt gctgatcccc 540
atgtgtatca ctggcaaaact gtgatggacg acaccgtcag tgcgtccgtc gcgcaggctc 600
tcgatgagct gatgctttgg gccgaggact gcccgaagt cccgcacctc gtgcacgcgg 660
atttcggctc caacaatgtc ctgacggaca atggccgcac aacagcggtc attgactgga 720
gcgaggcgat gttcggggat tcccaatacg aggtcgccaa catcttcttc tggaggccgt 780
ggttggcctt tatggagcag cagacgcgct acttcgagcg gaggcatccg gagcttgcat 840
gatcgccgcg gctccgggcg tatatgctcc gcattggtct tgaccaactc tatcagagct 900
tggttgacgg caatttcgat gatgcagctt gggcgaggg tcatgacgac gcaatcgtcc 960
gatccggagc cgggactgtc gggcgtagac aaatcgcccc cagaagcgcg gccgtctgga 1020
ccgatggctg tgtagaagta ctgcgccgata gtggaaaccg acgccccagc actcgtccga 1080
gggcaaagga atgcctgaga aaggaagtga gctgtaaagg ctgagctctc tctctgacgt 1140
atgtagcctc tggttagctt cgtcactcac tgttcttgac tcagcatggc aatctgatga 1200
aatcccagct gtaagtctgc ataaattgat gatctattaa acaataaaga tgtccactaa 1260
aatggaagtt ttttactgtc atactttgtt aagaaggggtg agaacagagt acctacattt 1320
tgaatggaag gattggagct acgggggtgg ggggtgggggt gggattagat aaatgcctgc 1380
tctttactga aggctcttta ctattgcttt atgataatgt ttcatagttg gatatacata 1440
tttaaacaag caaaaccaa ttaagggcca gctcattcct cactcatga tctatagatc 1500
tatagatctc atgtggatc tcgtggatc gaagcttatc gataccgtcg 1560
acctcgaggg ggggcccggg acccagcttt tgttcccttt agtgaggggt aattgcgcgc 1620
ttggcgtaat catggtcata gctgtttcct gtgtgaaatt gttatccgct cacaattcca 1680
cacaacatac gagccggaag cataaagtgt aaagcctggg gtgcctaata agtgagctaa 1740
ctcacattaa ttgcgttgcg ctcaactgcc gctttccagt cgggaaacct gtcgtgccag 1800
ctgcattaat gaatcgacca acgcgcgggg agaggcggtt tgcgtattgg gcgctcttcc 1860

```

gcttcctcgc	tactgactc	gctgcgctcg	gtcggttcggc	tgcggcgagc	ggatcagct	1920
cactcaaagg	cggtaatacg	gttatccaca	gaatcagggg	ataacgcagg	aaagaacatg	1980
tgagcaaaag	gccagcaaaa	ggccaggaac	cgtaaaaagg	ccgcgttgct	ggcgtttttc	2040
cataggctcc	gccccctga	cgagcatcac	aaaaatcgac	gctcaagtca	gaggtggcga	2100
aaccgcagac	gactataaag	ataccaggcg	tttccccctg	gaagctccct	cgtagcgtct	2160
cctgttccga	ccctgccgct	taccggatac	ctgtccgctt	ttctcccttc	gggaagcgtg	2220
gcgctttctc	atagctcacg	ctgtaggat	ctcagttcgg	tgtaggctcg	tcgctccaag	2280
ctgggctgtg	tgcacgaacc	ccccgttcag	cccgcaccgt	gcgccttata	cggttaactat	2340
cgtcttgagt	ccaaccgggt	aagacacgac	ttatcgccac	tggcagcagc	cactggtaac	2400
aggattagca	gagcgaggta	tgtaggcggt	gctacagagt	tcttgaagtg	gtggcctaac	2460
tacggctaca	ctagaaggac	agtatttggg	atctgcgctc	tgctgaagcc	agttaccttc	2520
ggaaaaagag	ttggtagctc	ttgatccggc	aaacaaacca	ccgctggtag	cggtggtttt	2580
tttgtttgca	agcagcagat	tacgcgcaga	aaaaaaggat	ctcaagaaga	tcctttgatc	2640
ttttctacgg	ggcttgacgc	tcagtggaac	gaaaactcac	gttaagggat	tttgggtcatg	2700
agattatcaa	aaaggatctt	cacctagatc	cttttaaatt	aaaaatgaag	ttttaaatca	2760
atctaaagta	tatatgagta	aacttggctt	gacagttacc	aatgcttaat	cagtggagca	2820
cctatctcag	cgatctgtct	atttcgttca	tccatagttg	cctgactccc	cgtagctgag	2880
ataactacga	tacgggaggg	cttaccatct	ggccccagtg	ctgcaatgat	accgcgagac	2940
ccacgctcac	cggctccaga	tttatcagca	ataaaccagc	cagccggaag	ggccgagcgc	3000
agaagtgggt	ctgcaacttt	atccgcctcc	atccagttct	ttaattgttg	ccgggaagct	3060
agagtaagta	gttcgccagt	taatagtttg	cgcaacgttg	ttgccattgc	tacaggcatc	3120
gtggtgtcac	gctcgtcgtt	tggtatgggt	tcattcagct	ccgggttcca	acgatcaagg	3180
cgagttacat	gatcccccat	gttgtcaaaa	aaagcgggta	gctccttcgg	tcctccgata	3240
gttgtcagaa	gtaagtgggc	cgtagtggtt	tcactcatgg	ttatggcagc	actgcataat	3300
tctcttactg	tcatgccatc	cgtaagatgc	ttttctgtga	ctgggtgagta	ctcaaccaag	3360
tcattctgag	aatagtgtat	gcggcgaccg	agttgctctt	gcccggcgct	aatacgggat	3420
aataccgcgc	cacatagcag	aactttaaaa	gtgctcatca	ttggaaaacg	ttcttcgggg	3480
cgaaaactct	caaggatctt	accgctgttg	agatccagtt	cgatgtaacc	cactcgtgca	3540
cccaactgat	cttcagcatc	ttttactttc	accagcggtt	ctgggtgagc	aaaaaacagg	3600
aggcaaaatg	ccgcaaaaaa	gggaataagg	gcgacacgga	aatgttgaat	actcatactc	3660
ttcctttttc	aatattattg	aagcatttat	caggggttatt	gtctcatgag	cggatacata	3720
tttgaatgta	tttagaaaaa	taaacaaata	gggggttccgc	gcacatttcc	ccgaaaagtg	3780
ccacctaata	tgtaagcgtt	aatattttgt	taaaattcgc	gttaaatttt	tgttaaatac	3840
gctcattttt	taaccaatag	gccgaaatcg	gcaaaatccc	ttataaatca	aaagaataga	3900
ccgagatagg	gttgagtgtt	gttccagttt	ggaaacaagag	tccactatta	aagaacgtgg	3960
actccaacgt	caaagggcga	aaaaccgtct	atcagggcga	tggccacta	cgtgaaccat	4020
caccctaata	aagttttttg	gggtcgaggt	gccgtaaagc	actaaatcgg	aaccctaaag	4080
ggagcccccg	atttagagct	tgacgggggaa	agccggcgaa	cgtggcgaga	aagggaaggga	4140
agaaagcgaa	aggagcgggc	gctagggcgc	tggcaagtgt	agcggtcacg	ctgcgcgtaa	4200
ccaccacacc	cgccgcgctt	aatgcgccgc	tacagggcgc	gtcccattcg	ccattcaggc	4260
tgcgcaactg	ttgggaaggg	cgatcgggtg	gggcctcttc	gctattacgc	cagctggcga	4320
aaggggggatg	tgctgcaagg	cgattaaagt	gggtaacgcc	agggttttcc	cagtcacgac	4380
gttgtaaaac	gacggccagt	gagcgcgcgt	aatacgactc	actatagggc	gaattggagc	4440
tccaccgcgg	tgcggccggg	ccatgcaggc	cacgacatga	taagatacat	tgatgagttt	4500
ggacaaacca	caactagaat	gcagtgaata	aaatgcttta	tttgtgaaat	ttgtgatgct	4560
attgctttat	ttgtaaccat	tataagctgc	aataaacaag	ttcttgtaca	gtcgtctcat	4620
gccgagagtg	atcccggcgg	cggtcacgaa	ctccagcagg	accatgtgat	cgcgcttctc	4680
gttgggggtct	ttgctcaggg	cggactggta	gctcaggtag	tggttgtcgg	gcagcagcac	4740
ggggccgctcg	ccgatggggg	tgttctgctg	gtagtggctg	gcgagctgca	cgtgccgctc	4800
ctcgatgttg	tggcggtatc	tgaagttcac	cttgatgccg	ttcttctgct	tgtcggccat	4860
gatatagacg	ttgtggctgt	tgtagttgta	ctccagcttg	tgccccagga	tgttgccgtc	4920
ctccttgaag	tcgatgccct	tcagctcgat	gcgggttcacc	aggggtgtcgc	cctcgaactt	4980
cacctcggcg	cgggtcttgt	agttgccgtc	gtccttgaag	aagatgggtg	gtcctgggac	5040
gtagccttcg	ggcatggcgg	acttgaagaa	gtcgtgctgc	ttcatgtggg	cggggtagcg	5100
ggcgaagcac	tgcaggccgt	agccgaagg	ggtcacgagg	gtggggccagg	gcacggggcag	5160
cttgccgggtg	gtgcagatga	acttcagggt	cagcttgccg	taggtggcat	cgccctcgcc	5220
ctcgccggac	acgctgaact	tgtggccggt	tacgtcgccg	tccagctcga	ccaggatggg	5280
caccaccccg	gtgaacagct	cctcgccctt	gtccaccatg	gtggcggttag	cttgatttga	5340

cagtggctgg	gggttgcgcc	gccgggtttt	ataggaagcc	acagcggcca	ctcgagccat	5400
aaaaggcaac	tttaggaacg	gcgggggggtg	attggattcg	agtcgtttat	tcaccggcct	5460
tgccgcacag	tgacgacattt	ttttaccccc	tctccctcc	ttttgcgggg	gaaaaaaaaa	5520
aaaaaaaaaa	aaaaaggaga	agagaaaaaa	agcgaagcgag	agagaaagcg	agattgagga	5580
agaggatgaa	gagtttgccg	atgggtgctg	gttccgtagg	cccagatgga	caagaatagc	5640
ccccgccctt	gcggacagta	tcccattcag	tgactcagat	cagatcaagc	ggccgccagt	5700
gtgatggata	tctgcagaat	tccagcacac	tggcggccgt	tactagtgga	tccgtacaag	5760
acagttagct	agttggcaaa	accaaggaga	attctgtgac	aagggaaact	caggacaaat	5820
cctgctaggc	agttctggag	gggaagcgcc	ctgcgcctgc	agtactgacc	tgtgaccata	5880
gccagggcag	caatgcagga	gaaggcagaa	atgtogatgt	tcatattctg	caagttggag	5940
gagaattcaa	caatggaatc	aatccattcc	ccaaagccac	gaacgcattg	caacctgtgc	6000
aagaccaccc	cattgcaaaa	gatgagttta	ccctccactg	ggttggacct	gcaattaata	6060
ccaaagagag	agaggggaga	aaaagagaga	gagaaaagct	aaacaactaa	ttacttgaag	6120
agcaataaat	gagggattta	agagtcacct	aattactgaa	gagttaatta	aaatgtagat	6180
ccagtggacc	ttgaaagggg	ttaattttct	aacaatcaag	aacgccccct	atcccacctc	6240
cattccattc	atccaaggct	tctgggctcg	cttctctctg	tctttttttg	ttttctttct	6300
ttttcttctc	ttctccgact	tccattttct	attctgtctt	tttctctacc	ccaccctctg	6360
gtttcccttc	ctccctttct	tttcttttct	tgattttctt	cacagcctcc	ctggattgtc	6420
tcctccctcc	cttattacct	gtatgctaata	cgaaggacaa	acagttctaa	gaaagctgat	6480
tcaaaaagca	ggtcttggtc	ggctttgggc	aggtctgcga	agccagggat	cttctctgcc	6540
cagccccgga	tgactctccat	ggagccagtc	aggagatcat	agaattgctg	gatatgctgg	6600
gtgtctatct	cactcatttg	atagtcaggg	tccgcctgga	actggaattt	catttttaaaa	6660
agcacttaat	gaggttctct	aaaatatata	accctgaaa	ttgctaacc	cgtttctaata	6720
aggggagcca	ggtttttata	acaattaaac	ctctctctga	ccagtaagga	aattagatgt	6780
tccggggcaa	ttaattcaat	tagggatggg	gctacgaggt	cgctgcttta	aatatgtaaa	6840
ttgtcatttc	catacagact	aaatacagtg	ccatccagcc	ctgggaaacg	ttcactctta	6900
tctccacaaa	acaattgaca	tgtagtagt	catgctagtc	ccagagtggg	cctcgagctg	6960
agacgccctt	ttgcaaatct	taataaaaatt	gcagtcctta	ccaggtttgc	tcaacatact	7020
tcctccaatt	aaagcccgca	tattttttcac	actgtggagta	aaaggatctg	accacttgat	7080
ccccacccc	actggcatct	ttacaccctc	tccttccctt	ccctgacccc	cagacagccc	7140
ccacatcctt	cccagcacca	gtgccaggaa	ggatgaggat	taaagcatct	atctctatgg	7200
aggactggga	ataaaagtaga	ttgggaaggg	ggagagtttg	cacagcccct	ggatttcatg	7260
gtagaggcac	tggtgtatat	ggaagaggag	gggcgatggg	gtagtggggg	agtgggaatt	7320
cacacagctt	caaaccctgt	ggcgtgtccc	tctcctcctt	cctccatcct	gtcctcatgt	7380
ctatgcccc	agctacatgg	atttgtctga	agatctgaat	cttgaggccc	tggccagagc	7440
tgcgaggcat	atacagcctt	gcttgccctt	ctttaccccc	gttgaatctg	agagttaatg	7500
acggatgtgg	ggaggggtcc	tgcccatctg	tcggtttgtc	cacatgatat	cccccccgcc	7560
agcttcttac	cctggaatag	tccaggctgg	tcatagccgg	gttggaagtcg	acatggggcc	7620
tgacgagggc	actgatcaga	ctcacccggg	gcgaaggggg	agaggggctcc	tgtgggctct	7680
tcggtttcga	gggcaaacga	cctctccggc	cttttaaact	gtctgtgcga	accactgcaa	7740
aggaagagcc	ctgtagcgcc	cgcttttccg	agcccaggcc	cagctgctgc	ctcggtccct	7800
ccccggggaa	ggccgcagcc	gcggggcacc	aggetgagcg	gctgagggcc	ccagtgcctg	7860
taaagccttc	actgactaga	agcattaaaa	aatgcggggg	tattttatgt	cttccctcaa	7920
atgggtcgta	tagttaaagg	agagaagggc	ctggcggtt	tctctaggga	aggccgggca	7980
agcaggcagc	tgcagggtcc	tggaggccat	actgaggggg	agtcggagat	ccccagcacc	8040
gggaagtgga	acgtgatgct	ggagtatgag	cagtggtttc	ctaaaggcgc	aaactggagg	8100
gtcggcagct	cccctcagcc	taccttcttt	ggtaccgggc	gatacagtca	actgtctttg	8160
acctttgtta	ctactctctt	ccgatgatga	tgctgcactt	attctatgct	gtctcaatgt	8220
tagaggcata	tcagtctcca	ctgaagccaa	tctatctgtg	acggcatctt	tattcacatt	8280
atcttgtaca	aataatcctg	ttaacaatgc	ttttatatcc	tgtaaagaat	ccattttcaa	8340
aatcatgtca	aggtctttct	gaggaaaaat	cagtagaaat	agctgttcca	gtctttctag	8400
ccttgattcc	acttctgtca	gatgtgccct	agtcagcgga	gaccttttgg	ttttgggaga	8460
gtacgcacac	tcccagttgt	tcttcagaca	cttggcgcac	ttcggttttt	ctttggagca	8520
cttgagcttt	ttaagtcggc	aaatatcgca	tgcttggttcg	atagaagaca	gtagcttcat	8580
ggtggcaagc	ttaagtttaa	acgctagaac	tagtggatcc	ggataagcca	gtaagcagtg	8640
ggttctctag	ttagccagag	agctctgctt	atatagacct	cccaccgtac	acgcctaccg	8700
cccatttgcg	tcaatggggc	ggagttgtta	cgacattttg	gaaagtcctc	ttgatttttg	8760
tgccaaaaca	aactcccatt	gacgtcaatg	gggtggagac	ttggaaatcc	ccgtgagtca	8820

```

aaccgctatc caccgccatt gatgtactgc caaaaccgca tcaccatggt aatagcgatg 8880
actaatacgt agatgtactg ccaagtagga aagtcccata aggtcatgta ctgggcataa 8940
tgccaggcgg gccatttacc gtcattgacg tcaatagggg gcgtacttgg catatgatac 9000
acttgatgta ctgccaaagt ggcagtttac cgtaaatact ccaccattg acgtcaatgg 9060
aaagtcccta ttggcgttac tatgggaaca tacgtcatta ttgacgtcaa tgggcggggg 9120
tcgttggggc gtcagccagg cgggccattt accgtaagtt atgtaacgcg gaactccata 9180
tatgggctat gaactaatga ccccgtaatt gattactatt aataactagt caataatcaa 9240
tgtcaacccc gggctgcagg aattctaccg ggtaggggag gcgcttttcc caaggcagtc 9300
tgtagcatgc gcttttagcag ccccgctggc acttggcgca tcacaagtgg cctctggcct 9360
cgcacacatt ccacatccac cggtagcgcc aaccggctcc cttctttggg ggccccttcg 9420
cgccaccttc tactcctccc ctagtcagga agttcccccc cgccccgcag ctgcgcgtcg 9480
gcaggacgtg acaaatggaa gtagcacgtc tcaactagtct cgtgcagatg gacaagcacc 9540
gctgagcaat ggaagcgggt aggcctttgg ggcagcggcc aatagcagct tggctccttc 9600
gctttctggg ctgagaggct ggggaaggggt ggggtccgggg gcgggctcag gggcgggctc 9660
aggggcgggg cgggcgcgaa ggtcctccgg acccggcatt c 9701

```

<210> 74

<211> 2379

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic nucleotide  
sequence for GR region of homology

<400> 74

```

cgtacaagac agttagctag ttggcaaaac caaggagaat tctgtgacaa gggaaactca 60
ggacaaatcc tgctaggcag ttctggaggg gaagcgccct gcgcctgcag tactgacctg 120
tgaccatagc cagggcagca atgcaggaga aggcagaaat gtcgatgttc atattctgca 180
agttggagga gaattcaaca atggaatcaa tccattcccc aaagccacga acgcattgca 240
acctgtgcaa gaccacccca ttgcaaaaga tgagtttacc ctccactggg ttggacctgc 300
aattaatacc aaagagagag aggggagaaa aagagagaga gaaaagctaa acaactaatt 360
acttgaagag caataaatga gggatttaag agtcacctaa ttactgaaga gttaattaaa 420
atgtagatcc agtggacctt gaaaggggtt aatttcataa caatcaagaa cgccccctat 480
cccacctcca ttccattcat ccaaggcttc tgggctcgct tcttctcgtc tttttttgtt 540
ttctttcttt ttcttccttt ctccgacttc catttctat tctgtctttt tctctacccc 600
accctctggt ttcccttcct ccctttcttt tcccttcttg atttctctca cagcctccct 660
ggattgtctc ctccctccct tattacctgt atgctaatac aaggacaaac agttctaaga 720
aagctgatcc aaaaagcagg tcttggtcgg ctttgggcag gtctgcgaag ccagggatct 780
tctctgcccc gccccgatg atctccatgg agccagtcag gagatcatag aattgctgga 840
tatgctgggt gtcacttcca ctcatattgat agtcagggtt cgcttggaac tgggaatttca 900
ttttaaaaag cacttaataa ggttctctaa aatatataac ccgtgaaatt gctaaccccc 960
tttctaatag gggagccagg tttttataac aattaaacct ctctctgacc agtaaggaaa 1020
ttagatgttc cggggcaatt aattcaatta gggatggtgc tacgaggtcg ctgctttaaa 1080
tatgtaaatt gtcatttcca tacagactaa atacagtgcc atccagccct gggaacgctt 1140
cactcttatc tccacaaaac aattgacatg ttagtattca tgctagtccc agagtgggcc 1200
tcgagctgag acgccctttt gcaaacttta ataaaattgc agtccctacc aggtttgctc 1260
aacatacttc ctccaattaa agcccgcata tttttcacag ctggagtaaa aggatctgac 1320
cacttgatcc cccaccccac tggcatcttt acacctctc cttcccttcc ctgaccccca 1380
gacagcccc acatccttcc ccagcacagt gccaggaagg atgaggatta aagcatctat 1440
ctctatggag gactgggaat aaagtagatt ggggaagggg agagtgtgca cagcccctgg 1500
atttcaggt agaggcactg gtgtatatgg aagaggagg gcgatgggg agtggggtag 1560
tggaatttca cacagcttca aaccctgtgg cgtgtccctc tctccttcc tccatctgt 1620
cctcatgtct atgccccag ctacatggat ttgtctgaag atctgaatct tgaggccctg 1680
gccagagctg cgaggcatat acagccttgc ttgcctttct ttacccccgt tgaatctgag 1740
agttaatgac gtagtggggg aggggtcctg cccatctgtc ggtttgtcca catgatatcc 1800
ccccgccag cttcttacc tggaaatagtc caggctggtc atagccgggt tggagtcgac 1860

```

atggggccctg	acgagggcac	tgatcagact	caccgggggc	gaagggggag	agggctcctg	1920
tgggctcttc	ggtttcgagg	gcaaacgacc	tctccggcct	tttaaactgt	ctgtgcgaac	1980
cactgcaaag	gaagagccct	gttagcgccg	cttttccgag	cccaggccca	gctgctgcct	2040
cggtcctctc	ccggggaagg	ccgcagccgc	ggggcaccag	gctgagcggc	tgagggcccc	2100
agtgcttgta	aagccttcac	tgactagaag	cattaaaaaa	tgcgggggta	ttttatgtct	2160
tcctccaaat	gggtcgtata	gttaaaggag	agaagggcct	ggcggctttc	tctagggag	2220
gccgggcaag	caggcagctg	cagggtcctg	gaggccatac	tgagggggag	tcggagatcc	2280
ccagcaccgg	gaagtggaac	gtgatgctgg	agtatgagca	gtggtttcct	aaaggcgcaa	2340
actggagggt	cggcagctcc	cctcagccta	ccttctttg			2379

&lt;210&gt; 75

&lt;211&gt; 9597

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence: Synthetic  
pCDGal4-DBD-GR sequence

&lt;400&gt; 75

taacaggatt	atgtgtacaa	gataatgtga	ataaagatgc	cgtcacagat	agattggcct	60
cagtggagac	tgatatgcct	ctaacattga	gacagcatag	aataagtgcg	acatcatcat	120
cggaagagag	tagtaacaaa	gggtcaaagac	agttgactgt	atcgccgggt	accattcagc	180
aggccactac	aggagtctca	caagaaacct	ctgaaaatcc	tggttaacaaa	acaatagtct	240
ctgcaacgtt	accacaactc	acccctaccc	tggtgtcact	gttggagggt	attgaacctg	300
aagtgttata	tgcaggatat	gatagctctg	ttccagactc	aacttggagg	atcatgacta	360
cgctcaacat	gttaggagg	cggcaagtga	ttgcagcagt	gaaatgggca	aaggcaatac	420
caggtaagat	gcaaaacata	aaagagcaac	tatataaacc	tttgtgtttt	cttcagcaaa	480
aacacttttg	cttttatatc	atcgtgagcc	catggcttat	cttggtttctc	ttagtctctg	540
ggactatgaa	ggggagagtc	aggtgaatac	aggtgatagg	gagttttataa	taaaacattt	600
acattactcc	ctgcttttca	aatcattatg	cacaggatgg	taatttcaca	taggatgatg	660
taatatcaga	attcaagtta	caagactcac	tcaaaaactcc	ttttacactg	aagtttgggg	720
aaagaaaatg	tttttagtta	attccatttg	ttttccttca	ttgtgccact	tttaaaaatc	780
aggttggttg	taagattggt	aaacatcaag	tatgttgatt	gtcaaaaattt	gtactaaagt	840
agaatgattt	taacccttca	ctaaatgaaa	tgtacacat	tgaatgtaat	tttaaagata	900
attttaataa	aaagttaccc	tattggaatt	tggtgtggaa	tggcagagg	caatgttagt	960
gtcagctctg	actttaaaga	caggggaattg	acaagcctgt	gttcacgcaa	atagttagg	1020
agagagcaag	aaagtaacct	gacctcctgt	cactccttgt	ttattaagg	ggaaagagg	1080
gtgaatagca	gggc aaatgt	tttgcttaac	tcattgatta	atacctcaag	ccaagattct	1140
tttctgtttt	ttaaaaatcaa	tacataatag	ttgtacatat	ttactgtaca	tatttatatt	1200
taggggggtac	atgtaataat	ttaataaaaag	catacaacgt	gtaaggatca	aatcagagta	1260
actgggatat	ccatcacctc	aaacatttgt	ttgggggaaca	ttccaaatct	tctcttttag	1320
ctattttgaa	atataaagta	aattatttgt	aactatagtc	atcctgttgt	gctactgaac	1380
actaaaactt	atttcttcta	actgtatttt	tgcaccgcgt	aaccattccc	gcttcacccc	1440
catcaccact	atctttccc	gtcactggta	accgccaagc	caagaatttt	ggctatttta	1500
ctattttagtt	catgtttact	taagcagaca	gaggtgacaa	aactggcttt	tttttttttt	1560
ttttacatta	aaagctatta	aaaagcacct	agggggctgg	gtgcgatggc	tcacgcctgt	1620
aatcccagca	ctttgggaag	cccagggtggg	tggatcagtt	gaggtcagga	gttcgagacc	1680
agcctggcca	gcatagcaaa	accccatctc	tactaaaatt	acaaaaatta	gccgggcatg	1740
gtggtatgaa	tctgtattcc	tagctacttg	ggaggctggc	actgagaatc	acttgaaccc	1800
gggaggcgga	ggttgcagtg	agccgagatg	gcaccattgc	actccagcct	gggtgacaga	1860
gcaagacttt	gtctcaatta	aaaaaaaaaa	aaaaaaaaaa	aacacaagag	ggtttgtgag	1920
tcttaaagtg	tcagatgaca	gaagaaaact	gtgtctacct	agtattttaat	ttccattttc	1980
tgttaggggt	gcccttggtt	tgacagggct	aattgatctc	attgctcctt	ggcaattccc	2040
acagagatga	tcttctgaag	agtgttgcc	cataccttta	tttctcttaa	ttcaggtttc	2100
aggaacttac	acctggatga	ccaaatgacc	ctactgcagt	actcctggat	gtttcttatg	2160
gcatttgctc	tgggggtggag	atcatataga	caatcaagtg	caaacctgct	gtgttttgct	2220

cctgatctga	ttattaatga	gtaagttgta	tgtgtgtcat	tttccctgta	ttcatagggg	2280
atctttaacc	agctgatgtt	ttcctgattg	actgctattg	tgataattca	ggactgaaac	2340
aatcctacta	ggtatctagg	atctaggcaa	actggaaata	gagttatgag	tgcttggggc	2400
aggacaagtg	taatgtaaag	caaatgtaca	tgtggcatta	ttactgtccc	aggacatggt	2460
tgaggatatt	taacagcata	tctgaggtta	gtaaagtctg	tcgcaagcaa	caaggaatct	2520
tactgtgata	tcattttacat	aaccctattc	cagaaagaaa	aaggagcatg	gtaaaactca	2580
tgtggattca	gtggggacaa	ttgtagatga	ggatatctag	gctgatgggg	tgggacatat	2640
ggaccagac	acaagaggta	tctctttgca	tggaagggt	caccagtggt	ctgtggttta	2700
agaatatggg	aacaaatttg	ttttgtttaa	ctgagagaag	accaagcctt	taagatttta	2760
taaatcagct	attctcttat	cctctaagct	tattcctgtg	tctgcgaaat	acttcagggtg	2820
tccatttccc	cttacctcat	tgcagttgtt	tcctcactcg	ttttctccct	ccagtgtaac	2880
gttcatcatg	ttggctaattg	tttgcttcct	caagcacagt	ctgactgcat	cacatatctc	2940
cccagtacac	agattgtctt	cagtatcttc	ccactgaccc	tccagtacat	attctgcatg	3000
atttcagact	ttccagaatc	tgacctcact	tcctctccca	ttgttttcct	tcacacactc	3060
ttcattccca	tccatccttt	ccagcatact	cttagactct	tggtgttcac	atcaccagat	3120
acacagcaga	gttggtcacat	cctagttaact	ctcactttct	accttgattt	actacttttc	3180
gtacccttag	cttattgcta	ttagtacaat	gtaaacaggg	agttcacaca	cacatacccc	3240
tggtctaaga	agaataaaaa	atgaaggaga	tttctgtttg	tatagaaaac	agaagtcacc	3300
ttgactttta	ttgccaaaaa	gaggactgtt	caaactactg	catcacaatg	taacaagatt	3360
aggtagttgg	atccaatttt	aaattaaactg	gtaaatatat	ttagttttctg	gggaaactga	3420
agacattatt	actcatcata	atcctaccat	gctgtttaaa	aaataccatg	ttggcagtat	3480
ttgtttttta	gtcactttct	aatatgtaat	ttgaaggcat	ttaagtggaa	ttaaaagcat	3540
aaacagattt	gtatgaaaca	ccaacttatc	ctgggtttata	aaactaacct	aatttagggg	3600
ttttattatt	agggcattca	gatttagctt	taagcagtca	cagcaaaatc	taatcatgcc	3660
acatacatte	cttacataaa	gtgggattta	taattttttt	tcctcaacag	atttacatta	3720
gtttcatttt	cattaaggga	tatgtacttc	ctattcttgt	gttctcatgc	tgctgcctgc	3780
ggccgctcga	gtctagaggg	cccgtttaaa	cccgtgatc	agcctcgact	gtgccttcta	3840
gttccagctg	atctgttgtt	tgccctcttc	ccgtgccttc	cttgaccctg	gaagggtcca	3900
ctcccactgt	cctttcctaa	taaaatgagg	aaattgcatc	gcattgtctg	agtaggtgtc	3960
attctattct	gggggggtggg	gtggggcagg	acagcaaggg	ggaggatttg	gaagacaata	4020
gcaggcatgc	tggggatgcg	gtgggctcta	tggcttctga	ggcggaaaga	accagctggg	4080
gctctagggg	gtatccccac	gcgcctgtga	gcggcgcat	aagcgcgggc	ggtgtggtgg	4140
ttacgcgcag	cgtgaccgct	acacttgcca	gcgccttagc	gcccgcctct	ttcgctttct	4200
tcctttcctt	tctcgccacg	ttcgccggct	ttcccgtca	agctctaaat	cgggggctcc	4260
ctttagggtt	ccgatttagt	gctttacggc	acctcgaccc	caaaaaactt	gatttaggtg	4320
atggttcacg	tagtgggcca	tcgcccgtat	agacggtttt	tcgccccttg	acgttggagt	4380
ccacgttctt	taatagtggg	ctcttgttcc	aaactggaac	aacactcaac	cctatctcgg	4440
tctattcttt	tgattttataa	gggattttgc	cgatttcggc	ctattgggtta	aaaaatgagc	4500
tgattttaaca	aaaatttaac	gcgaattaat	tctgtggaat	gtgtgtcagt	taggggtgtg	4560
aaagtcacca	ggctccccag	caggcagaag	tatgcaaagc	atgcatctca	attagttagc	4620
aaccagtggt	ggaaagtccc	caggctcccc	agcaggcaga	agtatgcaaa	gcattgcatc	4680
caattagtca	gcaaccatag	tcccgcctct	aactccgccc	atcccgcctc	taactccgcc	4740
cagttccgcc	cattctccgc	cccatggctg	actaattttt	tttatttatg	cagaggccga	4800
ggccgcctct	gcctctgagc	tattccagaa	gtagttagga	ggcttttttg	gaggcctagg	4860
cttttgcaaa	aagctcccgg	gagcttggtat	atccattttc	ggatctgatc	agcacgtgat	4920
gaaaaagcct	gaactcaccg	cgacgtctgt	cgagaagttt	ctgatcgaaa	agttcgacag	4980
cgtctccgac	ctgatgcagc	tctcgagggt	cgaagaatct	cgtgctttca	gcttcgatgt	5040
aggagggcgt	ggatatgtcc	tgcgggtaaa	tagctgcgcc	gatggtttct	acaaagatcg	5100
ttatgtttat	cggcactttg	catcggccgc	gctcccgatt	ccggaagtgc	ttgacattgg	5160
ggaattcagc	gagagcctga	cctattgcat	ctcccgcctg	gcacaggggtg	tcacgttgca	5220
agacctgcct	gaaaccgaac	tgcccgtgtg	tctgcagccg	gtcgcggagg	ccatggatgc	5280
gatcgctgcg	gccgatctta	gccagacgag	cgggttcggc	ccattcggac	cgcaaggaat	5340
cggccaatc	actacatggc	gtgatttcat	actcgcgatt	gctgatcccc	atgtgtatca	5400
ctggcaaaact	gtgatggacg	acaccgtcag	tgcgtccgtc	gcgcaggctc	tcgatgagct	5460
gatgctttgg	gccgaggact	gccccgaagt	ccggcacctc	gtgcacgcgg	atttcggctc	5520
caacaatgtc	ctgacggaca	atggccgcat	aacagcggtc	attgactgga	gcgaggcgat	5580
gttcggggat	tccaataacg	aggtcgccaa	catcttcttc	tggaggccgt	ggttggcttg	5640
tatggagcag	cagacgcgct	acttcgagcg	gaggcatccg	gagcttgacg	gatcgccgcg	5700

gctccggg	tatatgctcc	gcattgggtct	tgaccaactc	tatcagagct	tgggtgacgg	5760
caatttcgat	gatgcagctt	gggcgcaggg	tcgatgcgac	gcaatcgctc	gatccggagc	5820
cgggactgtc	gggcgtacac	aaatcgccc	cagaagcgcg	gccgtctgga	ccgatggctg	5880
tgtagaagta	ctcgccgata	gtggaaaccg	acgccccagc	actcgtccga	gggcaaaggga	5940
atagcacgtg	ctacgagatt	tcgattccac	cgccgccttc	tatgaaaggt	tgggcttcgg	6000
aatcgttttc	cgggacgccc	gctggatgat	cctccagcgc	ggggatctca	tgctggagtt	6060
cttcgcccac	cccaacttgt	ttattgcagc	ttataatggg	tacaaataaa	gcaatagcat	6120
cacaaatttc	acaaataaa	catTTTTTt	actgcattct	agttgtgggt	tgtccaaact	6180
catcaatgta	tcttatcatg	tctgtatacc	gtcgacctct	agctagagct	tggcgtaatc	6240
atgggtcatag	ctgtttcctg	tgtgaaattg	ttatccgctc	acaattccac	acaacatacg	6300
agccggaagc	ataaagtgtg	aagcctgggg	tgccctaata	gtgagctaac	tcacattaat	6360
tgcggttgcc	tactgcccc	ctttccagtc	gggaaacctg	tcgtgccagc	tgcattaatg	6420
aatcggccaa	cgcgcgggga	gaggcggttt	gcgtattggg	cgctcttcct	cttcctcgct	6480
cactgactcg	ctgcgctcgg	tcgttcggct	gcggcgagcg	gtatcagctc	actcaaaggc	6540
ggtaatacgg	ttatccacag	aatcagggga	taacgcagga	aagaacatgt	gagcaaaagg	6600
ccagcaaaag	gccaggaacc	gtaaaaaagg	cgcgttgctg	gcgtttttcc	ataggctccg	6660
ccccctgac	gagcatcaca	aaaatcgacg	ctcaagtcag	aggtggcgaa	acccgacagg	6720
actataaaga	taccaggcgt	ttccccctgg	aaagctccct	gtgcgctctc	ctgttccgac	6780
cctgccgctt	accggatacc	tgtccgcctt	tctcccttcg	ggaagcgtag	cgctttctca	6840
tagctcacgc	tgtaggatat	tcagttcggg	gtaggtcggt	cgctccaagc	tgggctgtgt	6900
gcacgaaccc	cccgttcagc	ccgaccgctg	cgccttatcc	ggtaactatc	gtcttgagtc	6960
caaccgggta	agacacgact	tatcgccact	ggcagcagcc	actggtaaca	ggattagcag	7020
agcgagggtat	ctaggcgggtg	ctacagagtt	cttgaagtgg	tggcctaact	acggctacac	7080
tagaagaaca	gtatttggtg	tctgcgctct	gctgaagcca	gttaccttcg	gaaaaagagt	7140
tggtagctct	tgatccggca	aacaaaccac	cgctggtagc	ggtttttttg	tttgcaagca	7200
gcagattacg	cgcagaaaaa	aaggatctca	agaagatcct	ttgatctttt	ctacgggggtc	7260
tgacgctcag	tggaaacgaaa	actcacgtta	agggattttg	gtcatgagat	tatcaaaaag	7320
gatcttcacc	tagatccttt	taaattaaaa	ctgaagtttt	aaatcaatct	aaagtatata	7380
tgagtaaaact	tggcttgaca	gttaccatag	cttaatcagt	gaggcaccta	tctcagcgat	7440
ctgtctatatt	cgttcatcca	tagttgcctg	actccccgtc	gtgtagataa	ctacgatacg	7500
ggaggggctta	ccatctggcc	ccagtgtctg	aatgataccg	cgagaccac	gtcaccggc	7560
tccagattta	tcagcaataa	accagccagc	cgggaaggcc	gagcgcagaa	gtggctcctgc	7620
aactttatcc	gcctccatcc	agtctattaa	ttgttgccgg	gaagctagag	taagtagttc	7680
gccagttaat	agtttgcgca	acgttggttg	cattgctaca	ggcatcgtag	tgtcacgtct	7740
tgcgttttgg	atggcttcat	tcagctccgg	tccccaacga	tcaaggcgag	ttactgtatc	7800
ccccatgttg	tgcaaaaaag	cggttagctc	cttcggctct	ccgatcggtg	tcagaagtaa	7860
gttgggccga	gtgttatcac	tcatggttat	ggcagcactg	cataattctc	ttactgtcat	7920
gccatccgta	agatgctttt	ctgtgactgg	tgagtactca	accaagtcac	tctgagaata	7980
gtgtatgcgg	cgaccgagtt	gctcttgccc	ggcgtcaata	cgggataata	ccgcgccaca	8040
tagcagaact	ttaaaaagtgc	tcatcattgg	aaaacgttct	tcggggcgaa	aactctcaag	8100
gatcttaccg	ctgttgagat	ccagttcgat	gtaacccact	cgtgcaccca	actgatcttc	8160
agcatctttt	actttcacca	gcgtttctgg	gtgagcaaaa	acaggaaggc	aaaatgccgc	8220
aaaaaaggga	ataagggcga	cacggaaatg	ttgaatactc	atactcttcc	tttttcaata	8280
ttattgaagc	atttatcagg	gttattgtct	catgagcgga	tacatatttg	aatgtattta	8340
gaaaaataaa	caaatagggg	ttccgcgcac	atttccccga	aaagtgccac	ctgacgtcga	8400
cggatcggga	gatctcccga	tcccctatgg	tgcaactctc	gtacaatctg	ctctgatgcc	8460
gcatagttaa	gccagtatct	gctccctgct	tgtgtgttgg	aggtcgctga	gtagtgcgcg	8520
agcaaaattt	aagctacaac	aaggcaaggc	ttgaccgaca	attgcatgaa	gaatctgctt	8580
agggttaggc	gttttgcgct	gcttcgcgat	gtacgggcca	gatatacgcg	ttgacattga	8640
ttattgacta	gttatttaata	gtaatcaatt	acgggggtcat	tagttcatag	cccatatatg	8700
gagttccgcg	ttacataact	tacggtaaat	ggcccgcctg	gctgaccgcc	caacgacccc	8760
gcgccattga	cgtaataaat	gacgtatggt	cccatagtaa	cgccaatagg	gactttccat	8820
tgacgtcaat	gggtggagta	tttacggtaa	actgcccact	tggcagtaca	tcaagtgtat	8880
catatgccaa	gtacgcccc	tattgacgtc	aatgacggta	aatggcccg	ctggcattat	8940
gcccagtaca	tgaccttatg	ggactttcct	acttggcagt	acatctacgt	attagtcac	9000
gctattacca	tggatgatgc	gttttggcag	tacatcaatg	ggcgtggata	gcggtttgac	9060
tcacggggat	ttccaagtct	ccacccatt	gacgtcaatg	ggagtttgtt	ttggcaccaa	9120
aatcaacggg	actttccaaa	atgtcgtaac	aactccgccc	cattgacgca	aatgggcggt	9180

```

aggcgtgtac ggtgggaggt ctatataagc agagctctct ggctaactag agaaccact 9240
gcttactggc ttatcgaaat taatacgact cactataggg agaccaagc tggctagcgt 9300
ttaaacttaa gcttgccacc atgaagctac tgtcttctat cgaacaagca tgcgatattt 9360
gccgacttaa aaagctcaag tgctccaaag aaaaaccgaa gtgcgccaag tgtctgaaga 9420
acaactggga gtgtcgctac tctcccaaaa ccaaaagggt tccgctgact agggcacatc 9480
tgacagaagt ggaatcaagg ctagaaagac tggaaacagct atttctactg atttttcctc 9540
gagaagacct tgacatgatt ttgaaaatgg attctttaca ggatataaaa gcattgt 9597

```

<210> 76

<211> 10872

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic  
pKI-Gal4-DBD-GR sequence

<220>

<221> modified\_base

<222> (8885)..(8885)

<223> any nucleotide

<400> 76

```

cagttcgatg taaccactc gtgcacccaa ctgatcttca gcattcttta ctttcaccag 60
cgtttctggg tgagcaaaaa caggaaggca aaatgccgca aaaaaggga taagggcgac 120
acggaaatgt tgaatactca tactcttctt ttttcaatat tattgaagca tttatcaggg 180
ttattgtctc atgagcggat acatatattga atgtatttag aaaaataaac aaataggggg 240
tccgcgcaca tttccccgaa aagtgccacc taaattgtaa gcgttaatat tttgttaaaa 300
ttcgcgttaa atttttggtt aatcagctca ttttttaacc aataggccga aatcggaaca 360
atcccttata aatcaaaaga atagaccgag atagggttga gtgttggtcc agtttggaac 420
aagagtccac tattaaagaa cgtggactcc aacgtcaaaag ggcgaaaaac cgtctatcag 480
ggcgtaggcc cactacgtga accatcaccc taatcaagtt ttttgggggt gaggtgccgt 540
aaagcactaa atcggaaacc taaagggagc ccccgattta gagcttgacg gggaaagccg 600
gcgaacgtgg cgagaaagga agggaaagaa gcgaaggag cggcgctag ggcgctggca 660
agtgtagcgg tcacgctgcg cgtaaccacc acaccgccg cgcttaatgc gccgctacag 720
ggcgcgctcc attcgccatt caggctgctc aactgttggg aaggcgcatc ggtgctggcc 780
tcttcgctat tacgccagct ggcgaaaagg ggatgtgctg caaggcgatt aagttgggta 840
acgccagggt tttcccagtc acgacgttgt aaaacgacgg ccagtgaagc cgcgtaatac 900
gactcactat agggcgcaat ggagctccac cgcggtgctg cggggccatg caggccacga 960
catgataaga tacattgatg agtttggaca aaccacaact agaatgcagt gaaaaaaatg 1020
ctttatttgt gaaatttgtg atgctattgc tttatttgta accattataa gctgcaataa 1080
acaagttctt gtacagctcg tccatgccga gagtgatccc ggcggcggtc acgaactcca 1140
gcaggaccat gtgatcgctc ttctcgttgg ggtctttgct cagggcggac tggtagctca 1200
ggtagtgggt gtcgggcagc agcacggggc cgtcgccgat ggggggtgtt tgctggtagt 1260
ggtcggcgag ctgcacgctg ccgtcctcga tgttggtggc gatcttgaag ttcaccttga 1320
tgccgttctt ctgcttgtcg gccatgatag agacgttgtg gctgttgtag ttgtactcca 1380
gcttggtccc caggatgttg ccgtcctcct tgaagtcgat gcccttcagc tcgatcggtt 1440
tcaccagggt gtcgccctcg aacttcacct cggcgcggtt cttgtagttg ccgtcgtcct 1500
tgaagaagat ggtgcgctcc tggacgtagc cttcgggcat ggcggacttg aagaagtcgt 1560
gctgcttcat gtggtcgggg tagcgggcga agcactgcag gccgtagccg aaggtggtca 1620
cgagggtggg ccagggcacg ggcagcttgc cgggtgtgca gatgaacttc agggtcagct 1680
tgccgtagggt ggcatcgccc tcgccctcgc cggacacgct gaacttgtag ccgtttacgt 1740
cgccgtccag ctcgaccagg atgggcacca ccccggtgaa cagctcctcg cccttgctca 1800
ccatggtggc gttagcttga tttgacagtg gctgggggtt gcgccgccgg gttttatagg 1860
aagccacagc ggccactcga gccataaaaag gcaactttag gaacggcggg gggtagttgg 1920
attcgagtcg tttattcacc ggccttgccg cacagtgcag cattttttta cccctctctc 1980
ctccttttg cgggggaaaa aaaaaaaaaa aaaaaaaaaa ggagaagaga aaaaaagcga 2040

```



gcgagagaga	aagcgagatt	gaggaagagg	atgaagagtt	tggcgatggg	tgctgggtcc	2100
gtaggcccag	atggacaaga	atagcccccg	cccttgccga	cagtatccca	ttcagtgact	2160
cagatcagat	caagcggccg	caggcagcag	catgagaaca	caagaatagg	aagtacatat	2220
cccttaaatga	aaatgaaact	aatgtaaaatc	tgttgaggaa	aaaaaattat	aaatcccact	2280
ttatgtaagg	aatgtatgtg	gcatgattag	atthttgctgt	gactgcttaa	agctaaatct	2340
gaatgcccta	ataataaaaa	ccctaaatta	ggtttagtttt	ataaaccagg	ataagttggg	2400
gtttcataca	aatctgttta	tgctttttaat	tccacttaaa	tgctttcaaa	ttacatatta	2460
gaaagtgact	aaaaaacaaa	tactgccaac	atgggtattht	ttaaacagca	tggtaggatt	2520
atgatgagta	ataatgtctt	cagtttcccc	agaaactaaa	tatatthtacc	agthtaattta	2580
aaattggatc	caactacctc	atcttggttac	attgtgatgc	agtagtttga	acagtcctct	2640
ttttggcaat	aaaagtcaag	gtgacttctg	ttttctatac	aaacagaaat	ctccttcatt	2700
ttttattctt	cttagaccag	gggtatgtgt	gtgtgaactc	cctgtttaca	ttgtactaat	2760
agcaataagc	taggggtacg	aaaagtagta	atacaaggta	gaaagtgaga	gtaactagga	2820
tgtgacttct	ctgctgtgta	tctgggtgatg	tgaacaccaa	gagtctaaga	gtatgctgga	2880
aaggatggat	gggaatgaag	agtgtgtgaa	ggaaaacaat	gggagaggaa	gtgaggtcag	2940
attctggaaa	gtctgaaatc	atgcagaata	tgtactggag	ggtcagtggg	aagatactga	3000
agacaatctg	tgtactgggg	agatatgtga	tgcagtcaga	ctgtgcttga	ggaagcaaac	3060
attagccaac	atgatgaacg	ttacactgga	gggagaaaaa	gagtgaggaa	acaactgcaa	3120
tgaggtaagg	ggaaatggac	acctgaagta	tttcgcagac	acaggaataa	gcttagagga	3180
taagagaata	gctgatthtat	aaaatcttaa	aggcttggtc	ttctctcagt	taaacaaaaa	3240
aaatttgtht	ccatattctt	aaaccacaga	cactgggtga	gccttgccat	gcaaagagat	3300
acctctgtg	tctgggtcca	tatgtcccac	cccatcgcc	tagatatcct	catctacaat	3360
tgtccccact	gaatccacat	gagthttacc	atgctccttt	ttctttctgg	aatagggtta	3420
tgtaaatgat	atcacagtaa	gattccttgt	tgcttgccag	agactthtact	aacctcagat	3480
atgctgttaa	atatcctcaa	acatgtcctg	ggacagtaat	aatgccacat	gtacatttgc	3540
tttacattac	acttgtcctg	ccccaaagcac	tcataactct	atthtccagtt	tgcttagatc	3600
ctagatacct	agtaggattg	thttcagthct	gaattatcac	aatagcagtc	aatcaggaaa	3660
acatcagctg	gttaaagata	ccctatgaat	acagggaaaa	tgacacacat	acaacttact	3720
cattaataat	cagatcagga	gcaaaaacaca	gcaggthtgc	acttgattgt	ctatatcact	3780
tccacccag	agcaaatgcc	ataagaaaca	tccaggagta	ctgcagtagg	gtcatttggt	3840
catccaggtg	taagthtctg	aaacctgaat	taagagaaat	aaagggtatga	ggcaacactc	3900
ttcagaagat	catctctgtg	ggaattgccca	aggagcaatg	agatcaatta	gccctgtcaa	3960
aacaagggca	cccctaacag	aaaatggaaa	ttaaatacta	ggtagacaca	gtthttcttct	4020
gtcatctgac	actthtaagac	tcacaaaacct	thttgtgttt	thttthttttt	thttthttttt	4080
aatthgagaca	aagthcttgct	ctgtcaccca	ggctggagtg	caatgggtgcc	atctcggctc	4140
actgcaacct	ccgcctcccg	ggthcaagtg	atthtccagtg	ccagcctccc	aagtagctag	4200
gaatacagat	tcataccacc	atgcccggct	aatthtttgta	atthtttagtag	agatgggggt	4260
ttgctatgct	ggccaggctg	gtctcgaact	cctgacctca	actgatccac	ccacctgggc	4320
thcccaaagt	gctgggatta	caggcgtgag	ccatcgcacc	cagcccccta	gggtgctthtt	4380
aatagcttht	aatgtaaaaa	aaaaaaaaaa	aagccagtht	tgthcactct	gtctgcttaa	4440
gtaaacatga	actaaatagt	aaaatagcca	aatthcttgg	cttggcggtt	accagtgacc	4500
gggaaagata	gtgggtgatgg	ggatgaagcg	ggaatgggtg	acgggtgcaa	aaatacagtt	4560
agaagaaata	agthtttagtg	thcagtagca	caacaggatg	actatagtta	acaataatth	4620
actthtatatt	tcaaaatagc	taaaagagaa	gatttggaat	gtthccccaa	caaagtgttg	4680
aggtgatgga	tatcccagtt	actctgattt	gatccttaca	cgttgtagtc	thttattaaa	4740
thattacatg	taccccctaa	atataaatat	gtacagtaaa	tatgtacaac	tattatgtat	4800
tgattthtaaa	aaacagaaaa	gaatcttggc	thgaggtatt	aatcaatgag	thaaagcaaaa	4860
catttgccct	gctattcaca	cctctthtccc	ccttaataaa	acaaggatga	caggaggtca	4920
ggthacttht	thgtctctct	cctaactatt	tgctgtaaca	caggcttgth	aattccctgt	4980
ctthaaagth	agagctgaca	ctaacattga	cctctgccat	thcacaccaa	atthccaatag	5040
ggtaacttht	atthtaaaat	atctthtaaaa	thacattcaa	tgtgtagcat	thcatttagt	5100
gaagggthta	aatcattcta	ctthtagtaca	aatthtgaca	atcaacatac	thgatgtthta	5160
ccaatcttht	aaacaacctg	atththtaaaa	gtggcacaat	gaaggaaaaa	aaatggaatt	5220
aactaaaaac	atththcttht	ccccaaactth	agtgtaaaaa	gagthtttgag	tgagthctgt	5280
aacttgaatt	ctgatattac	atcatcctat	gtgaaattac	catcctgtgc	ataatgattt	5340
gaaaagcagg	gagtaatgta	aatgtthttat	tataaaactcc	ctatcacctg	tattcacctg	5400
actctcccct	tcatagthccc	cagaactaag	agaaacaaga	taagccatgg	gtcacgatg	5460
atataaaaagc	caaagtgtth	thgtctgaaga	aaacacaaaag	gtthtatatag	thgtctcttht	5520

atgttttgea	tcttacctgg	tattgccttt	gcccatttca	ctgctgcaat	cacttgccgc	5580
cctcctaaca	tggtgagcgt	agtcacgatc	ctccaagttg	agtctggaac	agagctatca	5640
tatcctgcat	ataacacttc	aggttcaata	acctccaaca	gtgacaccag	ggtaggggtg	5700
agttgtggta	acgttgcagg	aactattggt	ttgttaccag	gattttcaga	ggtttcttgt	5760
gagactcctg	tagtggcctg	ctgaatggta	cccggcgata	cagtcaactg	tctttgacct	5820
ttgttactac	tctcttccga	tgatgatgtc	gcacttattc	tatgctgtct	caatgttaga	5880
ggcatatcag	tctccactga	agccaatcta	tctgtgacgg	catctttatt	cacattatct	5940
tgtacaaata	atcctgttaa	caatgctttt	atatcctgta	aagaatccat	tttcaaaatc	6000
atgtcaagg	cttctcgagg	aaaaatcagt	agaaatagct	gttccagctc	ttctagcctt	6060
gattccactt	ctgtcagatg	tgccctagtc	agcgagagac	ttttgggttt	gggagagtag	6120
cgacactccc	agttgttctt	cagacacttg	gcgcacttcg	gtttttcttt	ggagcacttg	6180
agctttttta	gtcggcaa	atcgcatgct	tggtcgatag	aagacagtag	cttcatgggtg	6240
gcaagcttaa	gtttaaacgc	tagaactagt	ggatccggat	aagccagtaa	gcagtgggtt	6300
ctctagttag	ccagagagct	ctgcttatat	agacctccca	ccgtacacgc	ctaccgccca	6360
tttgcgctcaa	tggggcggag	ttgttacgac	attttggaaa	gtcccgttga	ttttgggtgc	6420
aaaacaaact	cccattgacg	tcaatggggg	ggagacttgg	aaatccccgt	gagtcacacc	6480
gctatccacg	cccattgatg	tactgccaaa	accgcatcac	catggtaata	gcgatgacta	6540
atacgtagat	gtactgccaa	gtaggaaaagt	cccataagg	catgtactgg	gcataatgcc	6600
aggcggggcca	tttaccgtca	ttgacgtcaa	tagggggcgt	acttggcata	tgatacactt	6660
gatgtactgc	caagtgggca	gtttaccgta	aatactccac	ccattgacgt	caatggaaag	6720
tccctatttg	cgttactatg	ggaacatacg	tcattattga	cgtcaatggg	cggggggtcgt	6780
tgggcgggtca	gccaggcggg	ccatttaccg	taagttatgt	aacgcggaac	tccatatatg	6840
ggctatgaac	taatgacccc	gtaattgatt	actattaata	actagtcaat	aatcaatgtc	6900
aaccccgggc	tgcaggaatt	ctaccgggta	ggggaggcgc	ttttcccaag	gcagtctgga	6960
gcatgcgctt	tagcagcccc	gctggcactt	ggcgcatcac	aagtggcctc	tggcctcgca	7020
cacattccac	atccaccggt	agcgccaacc	ggctcccttc	tttgggtggc	ccttcgcgcc	7080
accttctact	cctcccttag	tcaggaagtt	cccccccgcc	ccgcagctcg	cgtcgtgcag	7140
gacgtgacaa	atggaagtag	cacgtctcac	tagtctcgtg	cagatggaca	agcaccgctg	7200
agcaatggaa	cgtgtagagc	ctttggggcca	gcggccaata	gcagcttggc	tccttcgctt	7260
tctgggctca	gaggctggga	aggggtgggt	ccgggggcgg	gctcaggggc	gggctcaggg	7320
gcggggcggg	cgcgaaggtc	ctccggaccc	ggcattctgc	acgcttcaaa	agcgcacgtc	7380
tgcgcgctg	ttctcctctt	cctcatctcc	gggccttcga	cctgcatgaa	aaagcctgaa	7440
ctcaccgcga	cgtctgtcga	gaagtttctg	atcgaaaagt	tcgacagcgt	ctccgacctg	7500
atgcagctct	cggagggcga	agaatctcgt	gctttcagct	tcgatgtagg	agggcgtgga	7560
tatgtcctcg	gggtaaatag	ctgcgcgcat	ggtttctaca	aagatcgta	tgtttatcgg	7620
cactttgcat	cggccgcgct	cccgattccg	gaagtgtctg	acattgggga	attcagcgag	7680
agcctgacct	attgcatctc	ccgccgtgca	cagggtgtca	cgttgcaaga	cctgcctgaa	7740
accgaactgc	ccgctgttct	gcagccggtc	gcggaggcca	tggatgcgat	cgctgcggcc	7800
gatcttagcc	agacgagcgg	gttcggccca	ttcggaaccgc	aaggaatcgg	tcaatacact	7860
acatggcgtg	atttcatatg	cgcgatttgt	gatccccatg	tgtatcactg	gcaaactgtg	7920
atggacgaca	ccgtcagtgc	gtccgtcgcg	caggctctcg	atgagctgat	gctttggggc	7980
gaggactgcc	ccgaagtccg	gcacctcgtg	cacgcggatt	tcggctccaa	caatgtcctg	8040
acggacaatg	gccgcataac	agcggtcatt	gactggagcg	aggcgatgtt	cggggattcc	8100
caatacgagg	tcgccaacat	cttcttcttg	aggcctgggt	tggcttgat	ggagcagcag	8160
acgcgctact	tcgagcggag	gcatccggag	cttgacggat	cgccgcggct	ccgggcgtat	8220
atgctccgca	ttggtcttga	ccaactctat	cagagcttgg	ttgacggcaa	tttcgatgat	8280
gcagcttggg	cgcagggctg	atgacgacga	atcgtccgat	ccggagccgg	gactgtcggg	8340
cgtacacaaa	tcgcccgcag	aagcgcggcc	gtctggaccg	atggctgtgt	agaagtactc	8400
gccgatagtg	gaaaccgacg	ccccagcact	cgtccgaggg	caaaggaatg	cctgagaaag	8460
gaagtgaact	gtaaaaggctg	agctctctct	ctgacgtatg	tagcctctgg	ttagcttctg	8520
cactcactgt	tcttgactca	gcatggcaat	ctgatgaaat	cccagctgta	agtctgcata	8580
aattgatgat	ctattaaaca	ataaagatgt	ccactaaaat	ggaagttttt	tactgtcata	8640
ctttgttaag	aagggtgaga	acagagtacc	tacattttga	atggaaggat	tggagctacg	8700
gggggtgggg	tgggggtggg	attagataaa	tgccgtctct	ttactgaagg	ctctttacta	8760
ttgctttatg	ataatgtttc	atagttggat	atcataattt	aaacaagcaa	aaccaaatta	8820
agggccagct	cattcctcca	ctcatgatct	atagatctat	agatctctcg	tgggatcatt	8880
gtttntctga	tccactggaa	gcttatcgat	accgtcgacc	tcgagggggg	gcccgggtacc	8940
cagcttttgt	tccctttagt	gagggttaat	tgcgcgcttg	gcgtaatcat	ggtcatagct	9000

```

gtttcctgtg tgaaattggt atccgctcac aattccacac aacatacgag ccggaagcat 9060
aaagtgtaaa gcctgggggtg cctaagtgtg gagctaactc acattaattg cgttgcgctc 9120
actgcccgtc ttccagtcgg gaaacctgtc gtgccagctg cattaatgaa tcggccaacg 9180
cgcggggaga ggcggtttgc gtattgggcg ctcttcgct tcctcgctca ctgactcgct 9240
gcgctcggtc gttcggctgc ggcgagcggt atcagctcac tcaaaggcgg taatacgggt 9300
atccacagaa tcaggggata acgcaggaaa gaacatgtga gcaaaaggcc agcaaaaggc 9360
caggaaccgt aaaaaggccg cgttgctggc gtttttccat aggctccgcc cccctgacga 9420
gcatcacaaa aatcgacgct caagtacagag gtggcgaaac ccgacaggac tataaagata 9480
ccaggcggtt cccctggaa gctccctcgt gcgctctcct gttccgacct tgccgcttac 9540
cggatacctg tccgcctttc tcccttcggg aagcgtggcg ctttctcata gctcacgctg 9600
taggtatctc agttcgggtg aggtcggtcg ctccaagctg ggctgtgtgc acgaaccccc 9660
cgttcagccc gaccgctgcg ccttatccgg taactatcgt cttgagtcca acccggttaag 9720
acacgactta tcgccactgg cagcagccac tggtaacagg attagcagag cgaggtagtg 9780
aggcggtgct acagagttct tgaagtgggt gcctaactac ggctacacta gaaggacagt 9840
atctgggtatc tgcgctctgc tgaagccagt taccttcgga aaaagagttg gtagctcttg 9900
atccggcaaaa caaacaccg ctggtagcgg tggttttttt gtttgcaagc agcagattac 9960
gcgcagaaaa aaaggatctc aagaagatcc tttgatcttt tctacggggg ctgacgctca 10020
gtggaacgaa aactcacgtt aagggtttt ggtcatgaga ttatcaaaaa ggatcttcac 10080
ctagatcctt ttaaattaaa aatgaagttt taaatcaatc taaagtatat atgagtaaac 10140
ttggtctgac agttaccaat gcttaatcag tgaggcacct atctcagcga tctgtctatt 10200
tcgttcaccc atagtgcct gactccccgt cgtgtagata actacgatac gggagggtt 10260
accatctggc cccagtgcg caatgatacc gcgagacca cgctcaccgg ctccagattt 10320
atcagcaata aaccagccag ccggaaggcg cgagcgaga agtggtcctg caactttatc 10380
cgctccatc cagtctatta attgttgccg ggaagctaga gtaagtagtt cgccagttaa 10440
tagtttgccg aacgttggtt ccattgctac aggcacgtg gtgtcacgct cgctggttg 10500
tatggcttca ttcagctccg gttcccaacg atcaaggcga gttacatgat ccccatgtt 10560
gtgcaaaaaa gcggttagct ccttcgggtc tccgatacgt gtcagaagta agttggccgc 10620
agtgttatca ctcattggtt tggcagcact gcataattct cttactgtca tgccatcgt 10680
aagatgcttt tctgtgactg gtgagtactc aaccaagtca ttctgagaat agtgtatgcg 10740
gcgaccgagt tgctcttgcc cggcgtcaat acgggataat accgcgccac atagcagaac 10800
tttaaaagtg ctcattcatt gaaaacgttc ttcggggcga aaactctcaa ggatcttacc 10860
gctgttgaga tc 10872

```

<210> 77

<211> 3250

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic nucleotide sequence for MR region of homology sequence

<400> 77

```

tttgtggtgc ttaaaaaatg agcatgtggt gcctcctatg gtccagggtta ataccagact 60
ttgataaaagt tctatggtag aacatgtttc cctggcacct gtcctccctg cctgacagac 120
tcaagctctt cttttgttct ttattttatt tatttatttt ttaagatgga ttctcattct 180
gtcgcccagg ctggagttca gtgacatgat attggtcac tgcacccag gtttcaagcg 240
attctcctgc ctgaggctca caagtagctg gggttacagg tgcctgccag gacgccaac 300
taaattttgt atttttagta gagacagggt ttcattgatg tggcaaggct ggtctcaaat 360
tcctgacctc aagtgatctg cttgcctcag cctcccaaa gctgggatta caggcatggg 420
ccactgcacc aggccttctt ttgtttttta tcacctcccc catctccagg ggcatcatgt 480
tatttttctg cctgtggagt ttcactgggc aagctcctct tcaaagaatc agtcagtggg 540
gctttcccca acttttacag ctgctgtctg accacaggta atctaagaaa ctgctttaga 600
ttaatgaact gggacaatgt aatggcaatt tctaagtgat gtttctgtag attaagaggg 660
gccttattaa ttaaatacca tgaaacaaaa aaaaatgact aaaataggaa ttccaggggac 720
aatttaaagg gtgctggggg ccagggtgatt attttttaga cagaagagtc attccaggtc 780
tcttacaact tacttgctag gccttgctt acttgggtta cacttagatt aaccctcgtt 840

```

```

atgagccgct aataatgtgc tcaacttggtg atttcatgaa cagaatgtat tttaccttat 900
taatatTTTT ggatgaaatt ttccatttttg atcaacataa gaagtgtatg tcattttttt 960
tatttttaatt ttcttttttt tttttgagac agagtctcac tctgttgctc aggcaggagt 1020
atagtgggtgg gatctcggct cactgcagcc tctgccgcc aggttcaagc gattctcctg 1080
cctcagcctc ctgagtagct gggattacag gcacctgcca ccaagcagga ctaatttttg 1140
tgtttttagt agaaacgggg tttcaccatc ttggctaggc tggctctgaa ctctgacct 1200
cgtgatccat ctgcctcagt aagccacctg ccttggcctc ccaaagtgtg gggattacag 1260
gcgtgagcca ctgcgccag ccgatatctgt cattttaaca ataaaatcct taaccctgca 1320
ttctcggaag aaatgagcac attaacata tttttcctat acacttgatc aatttgttat 1380
caatttaaaa aatctgatat caggatttca tagaacgcaa actcctcctg catggttggg 1440
gatggcgaa tcaagtgtccc agacttggtt agttcaggat gcagcctgtg aaaggagagg 1500
caatcctacc tggaagtacc tttgccact tcacgacttg gatcatctgt ttgcctgcta 1560
agcgggttag cgtggagagc agattttcgg ctgtatctgg ttttgagctg tcatagcctg 1620
catatacaat ttcaggttca atgttttcaa ggaccataac gggggaaggt gtgagcgctc 1680
gtgagattgt ggagagctga ggaaccagtg ctgtgttgac cgagggttct tttgcaggag 1740
cgatgtacgt tgtcccttcc tctgggcttt gcgggggtgg ggggtgggggt gggggctgct 1800
gctgctgtgg ctgctcctcg tgaatccctt tcttttcata ttttgatta cttttcttgc 1860
tttcatcacc aacctatttc tcatctgcct cctcccctac tgaaaccctt gttttcctat 1920
tgccagagag agccatctgt cagaatatgg ctgtagttac tatcataatt caaatattcc 1980
cttttaagta tttcttgacc caacaattac aaagaagcac tgctagatac aaacataaat 2040
aaaatgtgct aatgaggaat gcaaacagta ttaaaatgta aacttaaaat aggaaggaat 2100
ttgtggccaa agaaaaactg attacgattc tttcagttct ttaaaatgca tcataaaact 2160
aaaaatgttt tgctacagaa taacagtttg gcttctttta tgatggcaaa gaagggtaat 2220
tacgattgtc atgaaaacac attgagcagc tgtatactgt taacaaaaga accattatat 2280
gagctcgtaa aagctggcac tttatatatt cacatttctt agaaattaga ttgtcttccc 2340
aattggagtc gataccaaaa gagactgttt tctattaatt gtggctttgt accagtcaat 2400
tttcagagcc aactggatca tctttctttt gggatatatc ccacatatta ttatgtggaa 2460
tttatatata ctgatttata tacactcaga agcatacagt ataatgttgc aaagagtaat 2520
tttttaagtt cccagagatc tgtaaatttt taacgttaaa aaatgccagt ttcagtagat 2580
tctttcacca acgtacatct gttagaaat tactgaaaga aatcataacg cataactctg 2640
cagtttattt tatgaaggct aattaatagg tacttactca ttaaagacta ggtctggtgc 2700
aaaatagaga aattggctgt tcgtatgttt gtacgatctc cagctcaagg caaatgatga 2760
tagacacatc caagaatact ggattagggg aatttgggtc tcaagaggca agtttttaaa 2820
tcctgaagaa caaaacaatt aatcacagaa atacacttag catttaagta cattccagga 2880
agatgcttct taagtcaacc ccaaacagcc acctttcttt cacttttctc ttttctgaca 2940
gactcaaacc aattcaacag tcattttacag agtgcccacc actggcccag tcctgttcca 3000
catttcactt taaattcatc cttaaattcta gatgagggtg aactgccagg gcaaactaaa 3060
attgcaaagt tcagcaagtt tcagtgtgga tgacagccta aaacccttgg ttagattgag 3120
agaatttgta ttaacctgta aaacttgctt cgtcccaccc acaattaata aactaaaat 3180
aaagggtaat ttaataactg tcaaaccctg tgtatctgca ataatggtgg gagatttcat 3240
tgttcatgag

```

<210> 78

<211> 9212

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic  
pCDGal4-DBD-MR sequence

<400> 78

```

agcactgcat aattctctta ctgtcatgcc atccgtaaga tgcttttctg tgactggtga 60
gtactcaacc aagtcattct gagaatagtg tatgcggcga ccgagttgct cttgccgggc 120
gtcaatacgg gataataccg cgccacatag cagaacttta aaagtgtcga tcattggaaa 180
acgttcttcg gggcgaaaac tctcaaggat cttaccgctg ttgagatcca gttcgatgta 240
accactcgt gcaccaact gatcttcagc atcttttact ttcaccagcg tttctgggtg 300

```

agcaaaaaca	ggaaggcaaa	atgccgcaaa	aaaggggaata	agggcgacac	ggaaatgttg	360
aatactcata	ctcttccttt	ttcaatatta	ttgaagcatt	tatcagggtt	attgtctcat	420
gagcggatac	atatttgaat	gtatttagaa	aaataaacia	ataggggttc	cgcgacatt	480
tccccgaaaa	gtgccacctg	acgtcgacgg	atcgggagat	ctcccgatcc	cctatgggtg	540
actctcagta	caatctgctc	tgatgccgca	tagttaagcc	agtatctgct	ccctgcttgt	600
gtgttggagg	tcgctgagta	gtgcgcgagc	aaaatttaag	ctacaacaag	gcaaggcttg	660
accgacaatt	gcatgaagaa	tctgcttagg	gttaggcgtt	ttgcgctgct	tcgcgatgta	720
cgggccagat	atacgcgttg	acattgatta	ttgactagtt	attaatagta	atcaattacg	780
gggtcattag	ttcatagccc	atatatggag	ttccgcgtta	cataacttac	ggtaaattgg	840
ccgcctggct	gaccgcccac	cgacccccgc	ccattgacgt	caataatgac	gtatgttccc	900
atagtaacgc	caatagggac	tttccattga	cgtcaatggg	tggagtattt	acggtaaact	960
gcccacttgg	cagtacatca	agtgtatcat	atgccaaagta	cgccccctat	tgacgtcaat	1020
gacggtaaat	ggcccgccctg	gcattatgcc	cagtacatga	ccttatggga	ctttcctact	1080
tggcagta	tctacgtatt	agtcacgtct	attaccatgg	tgatgcgggt	ttggcagtac	1140
atcaatgggc	gtggatagcg	gtttgactca	cggggatttc	caagtctcca	ccccattgac	1200
gtccatggga	gtttgttttg	gcacccaaaat	caacgggact	ttccaaaatg	tcgtatgttc	1260
tccgccccat	tgacgcaaat	gggcggtagg	cgtgtacggg	gggagggtcta	tataagcaga	1320
gctctctggc	taactagaga	acccactgct	tactggctta	tcgaaattaa	tacgactcac	1380
tatagggaga	cccaagctgg	ctagcgttta	aacttaagct	tgccaccatg	aagctactgt	1440
cttctatcga	acaagcatgc	gatatttgcc	gacttaaaaa	gctcaagtgc	tccaaagaaa	1500
aaccgaagtg	cgccaagtgt	ctgaagaaca	actgggagtg	tcgctactct	cccaaaacca	1560
aaaggtctcc	gctgactagg	gcacatctga	cagaagtgga	atcaaggcta	gaaagactgg	1620
aacagctatt	tctactgatt	tttcctcgag	aagaccttga	catgattttg	aaaatggatt	1680
ctttacagga	tataaaagca	ttgttaacag	gattattttg	acaagataat	gtgaataaag	1740
atgccgtcac	agatagattg	gcttcagtgg	agactgatat	gcctctaaca	ttgagacagc	1800
atagaataag	tgcgacatca	tcatcggaag	agagtagtaa	caaagggtcaa	agacagttga	1860
ctgtatcgcc	gggtaccttt	gtggtgctta	aaaaatgagc	atgtgggtgcc	tcctatggtc	1920
caggtttaata	ccagactttg	ataaagtctt	atggtagaac	atgtttccct	ggcacctgtc	1980
ctccctgcct	gcagactca	agctcttctt	ttgtcttta	ttttatttat	ttatttttta	2040
agatggattc	tcattctgtc	gcccaggctg	gagttcagtg	acatgatatt	ggctcactgc	2100
atcccagggt	tcaagcgatt	ctcctgcctc	aggctcacia	gtagctgggg	ttacagggtg	2160
ctgccaggac	gcccactaa	attttgtatt	tttagtagag	acagggtttc	atgatgggtg	2220
caaggctggt	ctcaaattcc	tgacctcaag	tgatctgctt	gcctcagcct	cccaaagtct	2280
gggattacag	gcatgggcca	ctgcaccagg	ccttcttttg	ttttttatca	cctccccctc	2340
ctccaggggc	atcatgttat	tttctgcct	ctggagtttc	actgggcaag	ctcctcttca	2400
agaatcag	cagtggagct	ttccccaact	tttacagctg	ctgtctgacc	acaggtaatc	2460
taagaaactg	ctttagatta	atgaactggg	acaatgtaat	ggcaatttct	aagtgatgtt	2520
tctgtagatt	aagaggggcc	ttattaatta	aataccatga	aacaaaaaaa	aatgactaaa	2580
ataggaattc	cagggacaat	ttaaaggggtg	ctgggggcca	ggtgattatt	ttttagacag	2640
aagagtcatt	ccaggtctct	tacaacttac	ttgctaggcc	ttgccttact	tggtctacac	2700
ttagattaac	cctcgttatg	agccgctaata	aatgtgctca	acttgtgatt	tcataaacag	2760
aatgtatttt	acctatttaa	tatttttggg	tgaaattttc	catttttgatc	aacataagaa	2820
gtgtatgtca	ttttttttat	ttttaatttc	tttttttttt	ttgagacaga	gtctcactct	2880
gttgctcagg	ctggagtata	gtggtgggat	ctcggctcac	tgagcctct	gccgcccagg	2940
ttcaagcgat	tctcctgcct	cagcctcctg	agtagctggg	attacaggca	cctgccacca	3000
agcaggacta	atttttgtgt	tttttagtaga	aacgggggtt	caccatcttg	gtaggctggg	3060
tcttgaactc	ctgacctcgt	gatccatctg	cctcagtaag	ccacctgcct	tggcctccca	3120
aagtgcctggg	attacaggcg	tgagccactg	cgcccagccg	tatctgtcat	tttaacaata	3180
aaatccttaa	ccctgcattc	tcggaagaaa	tgagcacatt	aaccatattt	ttcctataca	3240
cttgatcaat	ttgttatcaa	tttaaaaaat	ctgatatcag	gatttcatag	aacgcaaact	3300
cctcctgcat	ggttggagat	ggcgaagtca	gttgcccaga	cttggttaagt	tcaggatgca	3360
gcctgtgaaa	ggagaggcaa	tcctacctgg	aagtaccttt	gcccacttca	cgacttggat	3420
catctgtttg	cctgctaagc	ggttgagcgt	ggagagcaga	ttttcggttg	tatctggttt	3480
tgagctgtca	tagcctgcat	atacaatttc	agggtcaatg	ttttcaagga	ccataacggg	3540
ggaagggtgtg	agcgctcgtg	agattgtgga	gagctgagga	accagtgcctg	tggtgaccga	3600
gggttctttt	gcaggagcga	tgtacgttgt	cccttcctct	gggctttgctg	ggggtggggg	3660
tgggggtggg	ggctgctgct	gctgtggctg	ctcctcgtga	atccctttct	tttcatcttt	3720
tgcattactt	ttcttgcttt	catcaccaac	ctattttctca	tctgcctcct	ccccactga	3780

aacccctggt	ttcctattgc	cagagagagc	catctgtcag	aatatggctg	tagttactat	3840
cataattcaa	atattccctt	ttaagtattt	cttgacccaa	caattacaaa	gaagcactgc	3900
tagatacaaa	cataaataaa	atgtgtcta	gaggaatgca	aacagtatta	aaatgtaa	3960
ttaaaatagg	aaggaatttg	tggccaaaga	aaaactgatt	acgattcttt	cagtctttta	4020
aaatgcatca	taaaactaaa	aatgttttgc	tacagaataa	cagtttggct	tcttttatga	4080
tggcaaagaa	gggtaattac	gattgtcatg	aaaacacatt	gagcagctgt	atactgttaa	4140
caaaagaacc	attatatgag	ctcgtaaaag	ctggcacttt	atatattcac	atttcctaga	4200
aattagattg	tcttcccaat	tggagtcgat	acccaaagag	actgttttct	attaattgtg	4260
gctttgtacc	agtcaatttt	cagagccaac	tggatcatct	ttcttttggg	atatatccca	4320
catattatta	tgtggaattt	atatatactg	atttatatac	actcagaagc	atacagtata	4380
atgttgcaaa	gagtaatttt	ttaagtcccc	agagatctgt	aaatttttta	cgtaaaaaaa	4440
tgccagtttc	agtagattct	ttcaccaacg	tacatctgtt	tagaaattac	tgaaagaaat	4500
cataacgcat	aactctgcag	tttattttat	gaaggcta	taataggtac	ttactcatta	4560
aagactaggt	ctgggtgcaaa	atagagaaa	tggctgttcg	tatgtttgta	cgatctccag	4620
ctcaaggcaa	atgatgatag	acacatccaa	gaatactgga	ttagggta	ttggctctca	4680
agaggcaagt	ttttaaatcc	tgaagaacaa	aacaattaat	cacagaaata	cacttagcat	4740
ttaagtacat	tccaggaaga	tgcctcttaa	gtcaacccca	aacagccacc	tttctttcac	4800
ttttctcttt	tctgacagac	tcaaataca	tcaacagtca	tttacagagt	gcccaccact	4860
ggcccagtc	tgttccacat	ttcactttta	attcatccta	aattctagat	gagggtaaac	4920
tgccagggca	aactaaaatt	gcaaagttca	gcaagtttca	gtgtggatga	cagcctaaaa	4980
cccttggtta	gattgagaga	atttgtatta	acctgtaaaa	cttgcttcgt	cccaccaca	5040
attaataaca	ctaaaataaa	gggtaattta	ataactgtca	aacctgtgt	atctgcaata	5100
atggtgggag	atttcattgt	tcatgaggcg	gccgctcgag	tctagagggc	ccgtttaaac	5160
ccgctgatca	gcctcgactg	tgccttctag	ttgccagcca	tctgttgttt	gccccctccc	5220
cgtgccttcc	ttgaccctgg	aagggtgccac	tcccactgtc	ctttccta	aaaatgagga	5280
aattgcatcg	cattgtctga	gtaggtgtca	ttctattctg	gggggtgggg	tggggcagga	5340
cagcaagggg	gaggattggg	aagacaatag	caggcatgct	ggggatgcgg	tgggctctat	5400
ggcttctgag	gcggaaagaa	ccagctgggg	ctctaggggg	tatccccacg	cgccctgtag	5460
cgccgcatcg	agcgcggcgg	gtgtgggtgg	tacgcgcagc	gtgaccgcta	cacttgccag	5520
cgccctagcg	cccgcctcct	tgcctttctt	cccttccctt	ctcgccacgt	tgcgcggctt	5580
tccccgctca	gctctaaatc	gggggctccc	tttagggttc	cgatttagtg	ctttacggca	5640
cctcgacccc	aaaaaacttg	attaggggtga	tggttcacgt	agtgggccat	cgccctgata	5700
gacggttttt	cgccctttga	cgttggagtc	cacgttcttt	aatagtggac	tcttgttcca	5760
aactggaaca	acactcaacc	ctatctcggt	ctattctttt	gatttataag	ggattttg	5820
gatttcggcc	tattgggttaa	aaaatgagct	gatttaacaa	aaatttaacg	cgaattta	5880
ctgtggaatg	tgtgtcagtt	aggggtgtga	aagtccccag	gctccccagc	aggcagaagt	5940
atgcaaagca	tgcattctca	ttagtcagca	accaggtgtg	gaaagtcccc	aggctcccca	6000
gcaggcagaa	gtatgcaaag	catgcatctc	aattagtcag	caaccatagt	cccgcacct	6060
actccgccca	tcccgcacct	aactccgccc	agttccgccc	attctccgcc	ccatggctga	6120
ctaatttttt	ttatttatgc	agaggccgag	gcccgcctct	cctctgagct	attccagaag	6180
tagtgaggag	gcttttttgg	aggcctaggg	ttttgcacaa	agctcccggg	agcttgtata	6240
tccatttttcg	gatctgatca	gcacgtgatg	aaaaagcctg	aactcaccgc	gacgtctgtc	6300
gagaagtttc	tgatcgaaaa	gttcgacagc	gtctccgacc	tgatgcagct	ctcggagggc	6360
gaagaatctc	gtgctttcag	cttcgatgta	ggagggcggt	gatatgtcct	gcgggtaaat	6420
agctgcgccg	atggtttcta	caaagatcgt	tatgtttatc	ggcactttgc	atcggccgcg	6480
ctcccgattc	cggaagtgtc	tgacattggg	gaattcagcg	agagcctgac	ctattgcac	6540
tcccgcctg	cacagggtgt	cacgttgcaa	gacctgcctg	aaaccgaact	gcccgtgtt	6600
ctgcagccgg	tgcgggaggg	catggatgcg	atcgctgcgg	ccgatcttag	ccagacgagc	6660
gggttcggcc	cattcggacc	gcaaggaa	ggtcaataca	ctacatggcg	tgatttcata	6720
tgcgcgattg	ctgatcccca	tgtgtatcac	tggcaaaactg	tgatggacga	caccgtcagt	6780
gcgtccgtcg	cgagggctct	cgatgagctg	atgctttggg	ccgaggactg	ccccgaagtc	6840
cggcacctcg	tgcacgcgga	tttcggctcc	aacaatgtcc	tgacggacaa	tggccgcata	6900
acagcggtca	ttgactggag	cgaggcgatg	tccggggatt	cccaatacga	ggtcgccaac	6960
atcttcttct	ggaggccgtg	gttggcttgt	atggagcagc	agacgcgcta	cttcgagcgg	7020
aggcatccgg	agcttgcagg	atcgccgcgg	ctccgggcgt	atatgtctcg	cattggctct	7080
gaccaactct	atcagagctt	ggttgacggc	aatttcgatg	atgcagcttg	ggcgagggtg	7140
cgatgcgacg	caatcgctccg	atccggagcc	gggactgtcg	ggcgtaacac	aatcgcccg	7200
agaagcgcg	ccgtctggac	cgatggctgt	gtagaagtac	tgcgcgatag	tggaaaccga	7260

```

cgccccagca ctcgtccgag ggcaaaggaa tagcacgtgc tacgagattt cgattccacc 7320
gccgccttct atgaaagggt gggcttcgga atcggtttcc gggacgccgg ctggatgac 7380
ctccagcgcg gggatctcat gctggagttc ttcccccacc ccaacttggt tattgcagct 7440
tataatgggt acaaataaag caatagcatc acaaatttca caaataaagc atttttttca 7500
ctgcattcta gttgtggttt gtccaaactc atcaatgtat cttatcatgt ctgtataccg 7560
tcgacctcta gctagagctt ggcgtaatca tgggtcatagc tgtttcctgt gtgaaattgt 7620
tatccgctca caattccaca caacatacga gccggaagca taaagtgtaa agcctgggg 7680
gcctaagtag tgagctaact cacattaatt gcgttgcgct cactgcccgc tttccagtcg 7740
ggaaacctgt cgtgccagct gcattaatga atcggccaac gcgcggggag aggcgggttg 7800
cgtattgggc gctcttcgcg ttctctcgct actgactcgc tgcgctcggg cgttcggctg 7860
cggcgagcgg tatcagctca ctcaaaggcg gtaatacggg tatccacaga atcaggggat 7920
aacgcaggaa agaacatgtg agcaaaaggc cagcaaaagg ccaggaaccg taaaaaggcc 7980
gcgttgctgg cgttttttcca taggctccgc cccctgacg agcatcaca aaatcgacgc 8040
tcaagtcaga ggtggcgaaa cccgacagga ctataaagat accaggcggt tccccctgga 8100
agctccctcg tgcgctctcc tgttccgacc ctgcccgtta ccggatacct gtccgccttt 8160
ctcccttcgg gaagcgtggc gctttctcat agctcacgct gtaggtatct cagttcgggtg 8220
taggtcgttc gctccaagct gggctgtgtg cacgaacccc ccgttcagcc cgaccgctgc 8280
gccttatccg gtaactatcg tcttgagtcc aaccgggtaa gacacgactt atcgccactg 8340
gcagcagcca ctggtaacag gattagcaga gcgaggtatg taggcgggtg tacagagttc 8400
ttgaagtggg ggcctaacta cggctacact agaagaacag tatttggtat ctgcgctctg 8460
ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt gatccggcaa acaaaccacc 8520
gctggtagcg gtttttttgt ttgcaagcag cagattacgc gcagaaaaaa aggatctcaa 8580
gaagatccct tgatcttttc tacgggggtct gacgctcagt ggaacgaaaa ctcacgttaa 8640
gggatttttg tcatgagatt atcaaaaagg atcttcacct agatcctttt aaattaaaaa 8700
tgaagtttta aatcaatcta aagtatatat gagtaaactt ggtctgacag ttaccaatgc 8760
ttaatcagtg aggcacctat ctacgctatc tgtctatttc gttcatccat agttgcctga 8820
ctccccgtcg tgtagataac tacgatacgg gagggcttac catctggccc cagtgtctga 8880
atgataccgc gagaccacg ctacccggct ccagatttat cagcaataaa ccagccagcc 8940
ggaagggcgg agcgcagaag tggctctgca actttatccg cctccatcca gtctattaat 9000
tggtgccggg aagctagagt aagtagttcg ccagttaata gtttgcgcaa cgttgttgcc 9060
attgctacag gcacgtgggt gtcacgctcg tcgtttggta tggcttcatt cagctccggg 9120
tccaacgat caaggcgagt tacatgatcc cccatgttgt gcaaaaaagc ggtagctcc 9180
ttcggctctc cgatcgttgt cagaagtaag tt 9212

```

<210> 79

<211> 10517

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic  
pKI-Gal-DBD-MR sequence

<220>

<221> modified\_base

<222> (9953)..(9953)

<223> any nucleotide

<400> 79

```

atcggacgat tgcgtcgcat cgaccctgcg cccaagctgc atcatcgaaa ttgccgtcaa 60
ccaagctctg atagagttgg tcaagacca tgcggagcat atacgcccg agccgcggcg 120
atcctgcaag ctccggatgc ctccgctcga agtagcgct ctgctgctcc atacaagcca 180
accacggcct ccagaagaag atgttggcga cctcgattg ggaatcccc aacatcgct 240
cgctccagtc aatgaccgct gttatgcggc cattgtccgt caggacattg ttggagccga 300
aatccgcgtg cacgaggtgc cggacttcgg ggcagtcctc ggcccaaagc atcagctcat 360
cgagagcctg cgcgacggac gcactgacgg tgcgtccat cacagtttgc cagtgatata 420
catggggatc agcaatcgcg catatgaaat cagcccatgt agtgtattga ccgattcctt 480

```

gcgggtccgaa	tggggccgaac	ccgctcgtct	ggctaagatc	ggccgcagcg	atcgcatcca	540
tggcctccgc	gaccggctgc	agaacagcgg	gcagttcggg	ttcaggcagg	tcttgcaacg	600
tgacaccctg	tgacggcg	gagatgcaat	aggtcaggct	ctcgctgaat	tccccaatgt	660
caagcacttc	cggaatcggg	agcgcggccg	atgcaaagt	ccgataaaca	taacgatctt	720
tgtagaaacc	atcggcgag	ctattttacc	gcaggacata	tccacgccct	cctacatcga	780
agctgaaagc	acgagattct	tcgccctccg	agagctgcat	caggctggag	acgctgtcga	840
acttttcgat	cagaaacttc	tcgacagacg	tcgcgggtgag	ttcaggcttt	ttcatgcagg	900
tcgaaggccc	ggagatgagg	aagaggagaa	cagcgcggca	gacgtgcgct	tttgaagcgt	960
gcagaatgcc	gggtccggag	gaccttcgcg	cccgccccgc	ccctgagccc	gcccctgagc	1020
ccgcccccg	acccaccctt	tcccagcctc	tgagcccaga	aagcgaagga	gccaagctgc	1080
tattggccgc	tgccccaaag	gcctaccgcg	ttccattgct	cagcgggtgct	tgtccatctg	1140
cacgagacta	gtgagacgtg	ctacttccat	ttgtcacgtc	ctgcacgacg	cgagctgcgg	1200
ggcggggggg	aacttcctga	ctaggggagg	agtagaagg	ggcgcgaagg	ggccaccaa	1260
gaaggggagcc	gggtggcgct	accgggtggat	gtggaatgtg	tgcgaggcca	gaggccactt	1320
gtgatgcgcc	aagtgccagc	ggggctgcta	aagcgcagtc	tccagactgc	cttgggaaaa	1380
gcgcctcccc	tacccggtag	aattcctgca	gcccgggggt	gacattgatt	attgactagt	1440
tattaatagt	aatcaattac	ggggtcatta	gttcatagcc	catatatgga	gttccgcgtt	1500
acataactta	cggtaaatgg	cccgcctggc	tgaccgcccc	acgacccccg	cccattgacg	1560
tcaataatga	cgtatgttcc	catagtaacg	ccaataggga	ctttccattg	acgtcaatgg	1620
gtggagtatt	tacggtaaac	tgcccacttg	gcagtacatc	aagtgtatca	tatgccaaat	1680
acgcccccta	ttgacgtcaa	tgacggtaaa	tggcccgccct	ggcattatgc	ccagtacatg	1740
accttatggg	actttcctac	ttggcagtag	atctacgtat	tagtcatcgc	tattaccatt	1800
gtgatgcggg	tttggcagta	catcaatggg	cgtggatagc	ggtttgactc	acggggattt	1860
ccaagtctcc	accccatgtg	cgtcaatggg	agtttggttt	ggcaccacaaa	tcaacgggac	1920
tttccaaaat	gtcgtaacaa	ctccgcccc	ttgacgcaaa	tgggcggtag	gcgtgtacgg	1980
tgggaggtct	atataagcag	agctctctgg	ctaactagag	aaccactgc	ttactggctt	2040
atccggatcc	actagttcta	gcgttttaac	ttaagcttgc	caccatgaag	ctactgtctt	2100
ctatcgaaca	agcatgcgat	atttgccgag	ttaaaaagct	caagtgcctc	aaagaaaaac	2160
cgaagtgcgc	caagtgtctg	aagaacaact	gggagtgtcg	ctactctccc	aaaaccacaaa	2220
gggtctccgct	gactagggca	catctgacag	aagtggaaatc	aaggctagaa	agactggaac	2280
agctattttct	actgattttt	cctcgagaag	accttgacat	gattttgaaa	atggattctt	2340
tacaggatat	aaaagcattg	ttaacaggat	tattttgtaca	agataatgtg	aataaagatg	2400
ccgtcacaga	tagattggct	tcagtggaga	ctgatatgcc	tctaacattg	agacagcata	2460
gaataagtcg	gacatcatca	tcggaagaga	gtagtaacaa	agggtcaaaga	cagttgactg	2520
tatcgccggg	tacctttgtg	gtgcttaaaa	aatgagcatg	tgggtgcctcc	tatggtccag	2580
gttaataacca	gactttgata	aagttctatg	gtagaacatg	tttccttggc	acctgtcttc	2640
cctgcctgac	agactcaagc	tcttcttttg	ttctttatatt	tattttattta	ttttttaaga	2700
tggattctca	ttctgtcgcc	caggctggag	ttcagtgaca	tgatattggc	tcactgcatc	2760
ccagggtttca	agcgattctc	ctgcctcagg	ctcacaagta	gctgggggtta	cagggtgcctg	2820
ccaggacgcc	caactaaatt	ttgtattttt	agtagagaca	gggtttcatg	atggtggcaa	2880
ggctgggtctc	aaattcctga	cctcaagtga	tctgcttgcc	tcagcctccc	aaagtctggg	2940
attacaggca	tgggccactg	caccaggcct	tcttttggtt	tttatcacct	cccccatctc	3000
caggggcatc	atgttatttt	tctgcctgtg	gagtttctact	gggcaagctc	ctcttcaaag	3060
aatcagtcag	tggagctttc	cccaactttt	acagctgctg	tctgaccaca	ggtaatctaa	3120
gaaactgctt	tagattaatg	aactgggaca	atgtaatggc	aatttctaag	tgatgtttct	3180
gtagattaag	aggggcctta	ttaattaaat	accatgaaac	aaaaaaaaat	gactaaaata	3240
ggaattccag	ggacaattta	aagggtgctg	ggggccagg	gattattttt	tagacagaag	3300
agtcattcca	gggtctcttac	aacttacttg	ctaggccttg	ccttacttgg	tctacactta	3360
gattaaccct	cgttatgagc	cgctaataat	gtgctcaact	tgtgatttca	tgaacagaat	3420
gtattttacc	ttattaatat	ttttggatga	aattttccat	tttgatcaac	ataagaagt	3480
tatgtcattt	tttttatttt	taattttctt	tttttttttg	agacagagtc	tcactctgtt	3540
gctcaggctg	gagtatagtg	gtgggatctc	ggctcactgc	agcctctgcc	gcccaggctt	3600
aagcgattct	cctgcctcag	cctcctgagt	acaggcgatt	acaggcacct	gccaccaagc	3660
aggactaatt	tttgtgtttt	tagtagaaac	ggggtttcac	catcttggct	aggctgggtc	3720
tgaactcctg	acctcgtgat	ccatctgcct	cagtaagcca	cctgccttgg	cctcccaaag	3780
tgctgggatt	acaggcgtga	gccactgcgc	ccagccgtat	ctgtcatttt	aacaataaaa	3840
tccttaaccc	tgcattctcg	gaagaaatga	gcacattaac	catatttttc	ctatacactt	3900
gatcaatttg	ttatcaattt	aaaaaatctg	atatcaggat	ttcatagaac	gcaaactcct	3960



cctgcatggt	tggagatggc	gaagtcagtt	gcccagactt	gttaagttca	ggatgcagcc	4020
tgtgaaagga	gaggcaatcc	tacctggaag	tacctttgcc	cacttcacga	cttggatcat	4080
ctgtttgcct	gctaagcggg	tgagcgtgga	gagcagattt	tcggctgtat	ctggttttga	4140
gctgtcatag	cctgcatata	caatttcagg	ttcaatggtt	tcaaggacca	taacggggga	4200
aggtgtgagc	gctcgtgaga	ttgtggagag	ctgaggaacc	agtgtgtgtg	tgaccgaggg	4260
ttctttttgca	ggagcgatgt	acgttgtccc	ttcctctggg	ctttgcgggg	gtgggggtgg	4320
gggtgggggc	tgctgctgct	gtggctgctc	ctcgtgaatc	cctttctttt	catcttttgc	4380
attacttttc	ttgctttcat	caccaaccta	tttctcatct	gcctcctccc	ctactgaaac	4440
ccctgttttc	ctattgccag	agagagccat	ctgtcagaat	atggctgtag	ttactatcat	4500
aattcaaata	ttccctttta	agtatttctt	gacccaacaa	ttacaaagaa	gcactgctag	4560
atacaaacat	aaataaaatg	tgctaattgag	gaatgcaaac	agtattaaaa	tgtaaaactta	4620
aaataggaag	gaatttgtgg	ccaaagaaaa	actgattacg	attcttttcag	tcttttaaaa	4680
tgcatcataa	aactaaaaat	gttttgctac	agaataacag	tttggttctt	tttatgatgg	4740
caaagaaggg	taattacgat	tgcatgaaa	acacattgag	cagctgtata	ctgttaacaa	4800
aagaaccatt	atatgagctc	gtaaaagctg	gcactttata	tattcacatt	tcctagaaat	4860
tagattgtct	tcccaattgg	agtcgatacc	aaaagagact	gttttctatt	aattgtggct	4920
ttgtaccagt	caattttcag	agccaactgg	atcatctttc	ttttgggata	tatcccacat	4980
attattatgt	ggaatttata	tatactgatt	tatatacact	cagaagcata	cagtataatg	5040
ttgcaaagag	taatttttta	agttcccaga	gatctgtaaa	tttttaacgt	taaaaaatgc	5100
cagtttcagt	agattctttc	accaacgtac	atctgttttag	aaattactga	aagaaatcat	5160
aacgcataac	tctgcagttt	atttttatgaa	ggctaattaa	taggtactta	ctcattaaag	5220
actaggctcg	gtgcaaaata	gagaaattgg	ctgttcgtat	gtttgtacga	tctccagctc	5280
aaggcaaatg	atgatagaca	catccaagaa	tactggatta	gggtaatttg	gtcctcaaga	5340
ggcaagtttt	taaatcctga	agaacaaaac	aattaatcac	agaaatacac	ttagcattta	5400
agtacattcc	aggaagatgc	ttcttaagtc	aaccccaaac	agccaccttt	ctttcacttt	5460
tctcttttct	gacagactca	aatcaattca	acagtcattt	acagagtgcc	caccactggc	5520
ccagtcctgt	tccacatttc	actttaaatt	catcctaaat	tctagatgag	ggtaaaactgc	5580
cagggcaaac	taaaattgca	aagttcagca	agttttcagt	tggatgacag	cctaaaacc	5640
ttggttagat	tgagagaatt	tgtattaaac	tgtaaaactt	gcttcgtccc	accacacatt	5700
aataacacta	aaataaaggg	taatttaata	actgtcaaac	cctgtgtatc	tgcaataatg	5760
gtgggagatt	tcattgttca	tgaggcggcc	gcttgatctg	atctgagtca	ctgaatggga	5820
tactgtccgc	aagggcgggg	gctattcttg	tccatctggg	cctacggaac	cagcacccat	5880
cgccaaactc	ttcatcctct	tcctcaatct	cgttttctct	ctcgtctgct	ttttttctct	5940
tctccttttt	tttttttttt	tttttttttt	ccccgcaaaa	ggaggggaga	gggggtaaaa	6000
aaatgctgca	ctgtgcggca	aggccgggtga	ataaacgact	cgaatccaat	cacccccgcg	6060
cgttcctaaa	gttgcccttt	atggctcgag	tggccgctgt	ggcttcctat	aaaaccgggc	6120
ggcgcaaccc	ccagccactg	tcaaatcaag	ctaaccgcac	catggtgagc	aagggcgagg	6180
agctgttcac	cggggtgggtg	cccatcctgg	tcgagctgga	cggcgacgta	aacggccaca	6240
agttcagcgt	gtccggcgag	ggcgagggcg	atgccacctc	cggcaagctg	accctgaagt	6300
tcatctgcac	caccggcaag	ctgcccgtgc	cctggcccac	cctcgtgacc	accttcgggt	6360
acggcctgca	gtgcttcgcc	cgctaccccc	accacatgaa	gcagcacgac	ttcttcaagt	6420
ccgccatgcc	cgaaggctac	gtccaggagc	gcaccatctt	cttcaaggac	gacggcaact	6480
acaagacccg	cgccgaggtg	aagttcgagg	gcgacaccct	ggtgaaccgc	atcgagctga	6540
agggcatcga	cttcaaggag	gacggcaaca	tcttggggca	caagctggag	tacaactaca	6600
acagccacaa	cgtctatatc	atggccgaca	agcagaagaa	cggcatcaag	gtgaacttca	6660
agatccgcca	caacatcgag	gacggcagcg	tgcagctcgc	cgaccactac	cagcagaaca	6720
ccccatcgg	cgacggcccc	gtgctgctgc	cgcgaacca	ctacctgagc	taccagtccg	6780
ccctgagcaa	agaccccaac	gagaagcgcg	atcacatggt	cctgctggag	ttcgtgaccg	6840
ccgccgggat	cactctcggc	atggacgagc	tgtacaagaa	cttgttttatt	gcagcttata	6900
atggttacaa	ataaagcaat	agcatcacia	atttcacaaa	taaagcattt	ttttcactgc	6960
attctagttg	tggtttgtcc	aaactcatca	atgtatctta	tcatgtcgtg	gcctgcatgg	7020
ccggcgcgca	ccgcgggtgga	gctccaattc	gccctatagt	gagtcgtatt	acgcgcgctc	7080
actggcgcg	cttttaaac	gtcgtgactg	ggcaaacctt	ggcgttacc	aacttaatcg	7140
ccttgacgca	catccccctt	tcgccagctg	gcgtaatagc	gaagaggccc	gcaccgatcg	7200
cccttcccaa	cagttgcgca	gcctgaatgg	cgaatgggac	gcgccctgta	gcggcgcat	7260
aagcgcggcg	ggtgtggtgg	ttacgcgcag	cgtgaccgct	acacttgcca	gcgccctagc	7320
gcccgcctct	ttcgctttct	tcccttccct	tctcgccacg	ttcgccggct	ttccccgtca	7380
agctctaaat	cgggggctcc	ctttaggggt	ccgatttagt	gctttacggc	acctcgaccc	7440

caaaaaactt	gattaggggtg	atgggttcacg	tagtggggcca	tcgccctgat	agacgggtttt	7500
tcgccctttg	acgttgggagt	ccacgtttctt	taatagtggg	ctcttgttcc	aaactggaac	7560
aacactcaac	cctatctcgg	tctattctttt	tgatttataa	gggattttgc	cgatttcggc	7620
ctattgggta	aaaaatgagc	tgattttaaca	aaaatttaac	gcgaatttta	acaaaatatt	7680
aacgcttaca	atttaggttg	cacttttcgg	ggaaatgtgc	gcggaacccc	tatttgttta	7740
tttttctaaa	tacattcaaa	tatgtatccg	ctcatgagac	aataaccctg	ataaatgctt	7800
caataatatt	gaaaaaggaa	gagtatgagt	attcaacatt	tccgtgtcgc	ccttattccc	7860
ttttttgcgg	cattttgcct	tcctgttttt	gctcacccag	aaacgctggg	gaaagtaaaa	7920
gatgctgaag	atcagttggg	tgacgagtg	ggttacatcg	aactggatct	caacagcggg	7980
aagatccttg	agagttttcg	ccccgaagaa	cgttttccaa	tgatgagcac	ttttaaaagt	8040
ctgctatgtg	gcgcggtatt	atccccgtatt	gacgcggggc	aagagcaact	cggtcgcgcg	8100
atacactatt	ctcagaatga	cttgggttgag	tactcaccag	tcacagaaaa	gcatcttacg	8160
gatggcatga	cagtaagaga	attatgcagt	gctgccataa	ccatgagtga	taacactgcg	8220
gccaaacttac	ttctgacaac	gatcggagga	ccgaaggagc	taaccgcttt	tttgacacaac	8280
atggggggatc	atgtaactcg	ccttgatcgt	tggaacccgg	agctgaatga	agccatacca	8340
aacgcgcgag	gtgacaccac	gatgcctgta	gcaatggcaa	caacgttgcg	caaactatta	8400
actggcgaac	tacttactct	agcttcccgg	caacaattaa	tagactggat	ggaggcggat	8460
aaagttgcag	gaccacttct	gcgctcgggc	cttcgggctg	gctggtttat	tgctgataaa	8520
tctggagccg	gtgagcgtgg	gtctcgcggt	atcattgcag	cactggggcc	agatggtaag	8580
ccctcccgtg	tcgtagtatt	ctacacgacg	gggagtcagg	caactatgga	tgaacgaaat	8640
agacagatcg	ctgagatagg	tgccctcactg	attaagcatt	ggtaactgtc	agaccaagtt	8700
tactcatata	tacttttagat	tgatttataa	cttcattttt	aattttaaaag	gatctagggtg	8760
aagatccttt	ttgataatct	catgacaaaa	atccccatac	gtgagttttc	gttccactga	8820
gcgtcagacc	ccgtagaaaa	gatcaaagga	tcttcttgag	atcctttttt	tctgcgcgta	8880
atctgctgct	tgcaaacaaa	aaaaccaccg	ctaccagcgg	tggtttgttt	gccggatcaa	8940
gagctaccaa	ctctttttcc	gaaggtaact	ggcttcagca	gagcgcagat	accaaatact	9000
gtccttctag	tgtagccgta	gttagggccac	cacttcaaga	actctgtagc	accgcctaca	9060
tacctcgctc	tgctaattcct	gttaccagtg	gctgtgcca	gtggcgataa	gtcgtgtctt	9120
accgggttgg	actcaagacg	atagttaccg	gataaggcgc	agcggtcggg	ctgaacgggg	9180
ggttcgtgca	cacagcccag	cttgagcgca	acgacctaca	ccgaactgag	atacctacag	9240
cgtgagctat	gagaaagcgc	cacgcttccc	gaagggagaa	aggcggacag	gtatccggta	9300
agcggcaggg	tcggaacagg	agagcgcacg	agggagcttc	cagggggaaa	cgcttggtat	9360
ctttatagtc	ctgtcgggtt	tcgccacctc	tgacttgagc	gtcgattttt	gtgatgctcg	9420
tcaggggggg	ggagcctatg	gaaaaacgcc	agcaacgcgg	ccttttttacg	gttcctggcc	9480
ttttgctggc	cttttgctca	catgttcttt	cctgcgttat	cccctgattc	tgtggataac	9540
cgtattaccg	cctttgagtg	agctgatacc	gctcgccgca	gccgaacgac	cgagcgcagc	9600
gagtcagtga	gcgaggaagc	ggaagagcgc	ccaatacgca	aaccgcctct	ccccgcgcgt	9660
tgggccgattc	attaatgcag	ctggcacgac	aggtttcccc	actggaaagc	gggcagtgag	9720
cgcaacgcaa	ttaatgtgag	ttagctcact	cattaggcac	cccaggcttt	acactttatg	9780
cttcgggctc	gtatgttggtg	tgggaattgtg	agcggataac	aatttcacac	aggaaacagc	9840
tatgaccatg	attacgccaa	gcgcgcgaatt	aacctcact	aaagggaaaca	aaagctgggt	9900
accgggcccc	ccctcgaggt	cgacggtatc	gataagcttc	cagtggatca	ganaaacaat	9960
gatcccacga	gagatctata	gatctataga	tcatgagtgg	aggaatgagc	tgggccttaa	10020
tttggttttg	cttggtttaa	ttatgatatc	caactatgaa	acattatcat	aaagcaatag	10080
taaagagcct	tcagtaaaga	gcaggcattt	atctaattcc	acccccaccc	ccacccccgt	10140
agctccaatc	cttccattca	aaatgtaggg	actctgttct	caccttctt	aacaaagtat	10200
gacagtaaaa	aacttccatt	ttagtggaca	tctttattgt	ttaatatagtc	atcaatttat	10260
gcagacttac	agctgggatt	tcatcagatt	gccatgctga	gtcaagaaca	gtgagtgcag	10320
aagctaacca	gaggctacat	acgtcagaga	gagagctcag	cctttacagc	tcacttcctt	10380
tctcaggcat	tcctttgccc	tcggacgagt	gctggggcgt	cggtttccac	tatcggcgag	10440
tacttctaca	cagccatcgg	tccagacggc	cgcgcttctg	cgggcgattt	gtgtacgccc	10500
gacagtcccg	gctccgg					10517

<210> 80  
 <211> 32  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: Synthetic primer

<400> 80  
 gatctagaga gccacacccc atgttgtctc ac 32

<210> 81  
 <211> 34  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: Synthetic primer

<400> 81  
 gagcggccgc ttgccaaggc ccctgtgaag cagg 34

<210> 82  
 <211> 5351  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: Synthetic pKI-CMV-SD sequence

<400> 82  
 tcccttttagt gagggttaat tgcgcgcttg gcgtaatcat ggtcatagct gtttcctgtg 60  
 tgaaattggt atccgctcac aattccacac aacatacgag ccggaagcat aaagtgtaaa 120  
 gcctgggggtg cctaattgagt gagctaactc acattaattg cgttgcgctc actgcccgtc 180  
 ttccagtcgg gaaacctgtc gtgccagctg cattaatgaa tcggccaacg cgcggggaga 240  
 ggcggtttgc gtattgggcg ctcttccgct tctcgtctca ctgactcgct gcgctcggtc 300  
 gttcggctgc ggcgagcggc atcagctcac tcaaaggcgg taatacgggt atccacagaa 360  
 tcaggggata acgcaggaaa gaacatgtga gcaaaaggcc agcaaaaggc caggaaccgt 420  
 aaaaaggccg cgttgctggc gtttttccat aggctccgcc cccctgacga gcatcacaaa 480  
 aatcgacgct caagtcagag gtggcgaaac ccgacaggac tataaagata ccaggcgttt 540  
 cccctggaa gctccctcgt gcgctctcct gttccgacct tgccgcttac cggatacctg 600  
 tccgcctttc tcccttcggg aagcgtggcg ctttctcata gctcacgctg taggtatctc 660  
 agttcgggtg aggtcgttcg ctccaagctg ggctgtgtgc acgaaccccc cgttcagccc 720  
 gaccgctgcg ccttatccgg taactatcgt cttgagtcca acccggtgaa acacgactta 780  
 tcgccactgg cagcagccac tggtaacagg attagcagag cgaggtatgt aggcggtgct 840  
 acagagttct tgaagtgggt gcctaactac ggctacacta gaaggacagt atttggtatc 900  
 tgcgctctgc tgaagccagt taccttcgga aaaagagttg gtagctcttg atccggcaaa 960  
 caaaccaccg ctggtagcgg tgggtttttt gtttgcaagc agcagattac gcgcagaaaa 1020  
 aaaggatctc aagaagatcc tttgatcttt tctacggggt ctgacgctca gtggaacgaa 1080  
 aactcacgtt aagggatttt ggtcatgaga ttatcaaaaa ggatcttcac ctagatcctt 1140  
 ttaaatataa aatgaagttt taaatcaatc taaagtatat atgagtaaac ttggtctgac 1200  
 agttaccaat gcttaatcag tgaggcacct atctcagcga tctgtctatt tcgttcatcc 1260  
 atagttgcct gactccccgt cgtgtagata actacgatac gggagggcct accatctggc 1320

cccagtgctg	caatgatacc	gcgagaccca	cgctcaccgg	ctccagattt	atcagcaata	1380
aaccagccag	ccggaagggc	cgagcgcaga	agtgggcctg	caactttatc	cgccctccatc	1440
cagtctatta	attgttgccg	ggaagctaga	gtaagtagtt	cgccagttaa	tagtttgccg	1500
aacggttggtg	ccattgctac	aggcatcggtg	gtgtcacgct	cgtcggtttgg	tatggcttca	1560
ttcagctccg	gttcccaacg	atcaaggcga	gttacatgat	cccccatggt	gtgcaaaaaa	1620
gcggttagct	ccttcggtcc	tccgatcggt	gtcagaagta	agttggccgc	agtgttatca	1680
ctcatggtta	tggcagcact	gcataattct	cttactgtca	tgccatccgt	aagatgcttt	1740
tctgtgactg	gtgagtactc	aaccaagtca	ttctgagaat	agtgtatgcg	gcgaccgagt	1800
tgctcttgcc	cggcgtaaat	acgggataat	accgcgccac	atagcagaac	tttaaaagtg	1860
ctcatcattg	gaaaacgttc	ttcggggcga	aaactctcaa	ggatcttacc	gctggtgaga	1920
tccagttcga	tgtaaccac	tcgtgcaccc	aactgatctt	cagcatcttt	tactttcacc	1980
agcgtttctg	ggtgagcaaa	aacaggaagg	caaaatgccg	caaaaaagg	aataagggcg	2040
acacggaaat	gttgaaatact	catactcttc	ctttttcaat	attattgaag	catttatcag	2100
ggttattgtc	tcatgagcgg	atacatattt	gaatgtattt	agaaaaataa	acaaataggg	2160
gttcgcgcga	catttccccg	aaaagtgcga	cctaaattgt	aagcgttaat	attttgtaa	2220
aattcgcgtt	aaatttttgt	taaatcagct	cattttttaa	ccaataggcc	gaaatcggca	2280
aaatccctta	taaatcaaaa	gaatagaccg	agatagggtt	gagtgttgtt	ccagtttgga	2340
acaagagtcc	actattaaag	aacgtggact	ccaacgtcaa	agggcgaaaa	accgtctatc	2400
agggcgatgg	cccactacgt	gaaccatcac	cctaatacag	ttttttgggg	tcgaggtgcc	2460
gtaaagcact	aaatcggaac	cctaaaggga	gcccccgatt	tagagcttga	cggggaaagc	2520
cggcgaaagt	ggcgagaaa	gaagggaaga	aagcgaaaag	agcgggcgct	agggcgctgg	2580
caagtcgtagc	ggtcacgtg	cgcgtaacca	ccacaccgcg	cgcgcttaat	gcgcgcgtac	2640
agggcgcgctc	ccattcgcca	ttcaggctgc	gcaactgttg	ggaaggcgga	tcggtgcggg	2700
cctcttcgct	attacgccag	ctggcgaaag	ggggatgtgc	tgcaaggcga	ttaagttggg	2760
taacgccagg	gttttcccag	tcacgacgtt	gtaaaacgac	ggccagtgag	cgcgcgtaat	2820
acgactcact	ataggcgcaa	ttggagctcc	accgcggtgg	cggccgctct	agaactagt	2880
gatcccccg	gaccggtccc	cccactcacc	tgccagtaag	cagtgggttc	tctagttagc	2940
cgagagcgtg	tgcttatata	gacctcccac	cgtcacgcgc	taccgcccac	ttgcgtcaat	3000
ggggcgaggt	tggttacgaca	ttttggaaa	tcccgttgat	tttgggtgcca	aaacaaactc	3060
ccattgacgt	caatgggggtg	gagacttgga	aatccccgtg	agtcaaaccg	ctatccacgc	3120
ccattgatgt	actgccaaaa	ccgcatcacc	atggtaatag	cgatgactaa	tacgtagatg	3180
tactgccaa	taggaaagt	ccataagggtc	atgtactggg	cataatgcca	ggcgggccc	3240
ttaccgtcat	tgacgtcaat	agggggcgta	cttggcatat	gatacacttg	atgtactgcc	3300
aagtggcg	tttaccgtaa	atagtccacc	cattgacgtc	aatggaaaagt	ccctattggc	3360
gttactatgg	gaacatacgt	cattattgac	gtcaatgggc	gggggtcggt	gggcgggtcag	3420
ccaggcgggc	catttaccgt	aagtattgta	acgcggaact	ccatatatgg	gctatgaact	3480
aatgacccc	taattgatta	ctattaataa	ctagtcaata	atcaatgtca	acgcccgggc	3540
tgcaaggaatt	ctaccgggta	ggggaggcgc	ttttcccaa	gcagtcctgga	gcagtcgctt	3600
tagcagcccc	gctggcactt	ggcgcatcac	aagtggcctc	tggcctcgca	cacattccac	3660
atccaccggt	agcgccaacc	ggctcccttc	tttgggtggc	ccttcgcgcc	accttctact	3720
cctcccctag	tcagggaagtt	cccccccgcc	ccgcagctcg	cgctcgtcag	gacgtgacaa	3780
atggaagtag	cacgtctcac	tagtctcggtg	cagatggaca	agcaccgctg	agcaatggaa	3840
gcgggtaggc	ctttggggca	gcggccaata	gcagcttggc	tccttcgctt	tctgggtcca	3900
gaggctggga	aggggtgggt	ccggggggcg	gctcaggggc	gggctcagg	gcggggcggg	3960
cgcaaggtc	ctccggaccc	ggcattctgc	acgcttcaaa	agcgcacgtc	tgccgcgctg	4020
ttctcctctt	cctcatctcc	gggccttcga	cctgcataaa	aaagcctgaa	ctcaccgcga	4080
cgtctgtcga	gaagtctctg	atcgaaaagt	tcgacagcgt	ctccgacctg	atgcagctct	4140
cggagggcga	agaatctcgt	gctttcagct	tcgatgtagg	agggcggtgga	tatgtcctgc	4200
gggtaaatag	ctgcgccgat	ggtttctaca	aagatcgtta	tgtttatcgg	cactttgcat	4260
cggccgcgct	cccgattccg	gaagtgcctg	acattgggga	attcagcgag	agcctgacct	4320
attgcatctc	ccgccgtgca	cagggtgtca	cgttgcaaga	cctgcctgaa	accgaactgc	4380
ccgctgttct	gcagccgggtc	gcggaggcca	tgagtgcat	cgctgcggcc	gatcttagcc	4440
agacgagcgg	gttcggccca	ttcggaccgc	aaggaatcgg	tcaatacact	acatggcgctg	4500
atctcatatg	cgcgattgct	gatccccatg	tgtatcactg	gcaaactgtg	atggacgaca	4560
ccgtcagtcg	gtccgtcgcg	caggctctcg	atgagctgat	gctttggggc	gaggactgcc	4620
ccgaagtccg	gcacctcggtg	cacgcggatt	tcggctccaa	caatgtcctg	acggacaatg	4680
gccgcataac	agcggtcatt	gactggagcg	aggcgatggt	cggggattcc	caatacgagg	4740
tcgccaacat	cttcttctg	aggccgtggt	tggcttgat	ggagcagcag	acgcgctact	4800

```

tcgagcggag gcatccggag cttgcaggat cgccgcggct cggggcgat atgctccgca 4860
ttggtcttga ccaactctat cagagcttgg ttgacggcaa tttcgatgat gcagcttggg 4920
cgcagggtcg atgcgacgca atcgctccgat ccggagccgg gactgtcggg cgtacacaaa 4980
tcgcccgcag aagcgcggcc gtctggaccg atggctgtgt agaagtactc gccgatagtg 5040
gaaaccgacg ccccgacact cgtccgaggg caaaggaatg agctcgctga tcagcctcga 5100
ctgtgccttc tagttgccag ccatctgttg tttgcccctc ccccgctgct tccattgtc 5160
tgagtaggtg tcattctatt ctggggggtg ggggtggggca ggacagcaag ggggaggatt 5220
gggaagacaa tagcaggcat gctggggatg cgggtgggctc tatggcttct gaggcggaaa 5280
gaaccagctg ggggatatca agcttatcga taccgtcgac ctcgaggggg ggcccggtag 5340
ccagcttttg t

```

&lt;210&gt; 83

&lt;211&gt; 9131

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence: Synthetic  
pKI-CMV-SD-Vanilloid sequence

&lt;400&gt; 83

```

cgatcgctgc ggccgatctt agccagacga gcgggttcgg cccattcgga ccgcaaggaa 60
tcgggtcaata cactacatgg cgtgatttca tatgcgcgat tgctgatccc catgtgtatc 120
actggcaaac tgtgatggac gacaccgtca gtgcgtccgt cgcgcaggct ctcgatgagc 180
tgatgctttg ggccgaggac tgccccgaag tccggcacct cgtgcacgcg gatttcggct 240
ccaacaatgt cctgacggac aatggccgca taacagcggg cattgactgg agcgaggcga 300
tgttcgggga ttcccaatac gaggtcgcca acatcttctt ctggaggccg tggttggtt 360
gtatggagca gcagacgcgc tacttcgagc ggaggcatcc ggagcttgca ggatcgccgc 420
ggctccgggc gtatatgctc cgcattggtc ttgaccaact ctatcagagc ttggttgacg 480
gcaatttcga tgatgcagct tgggcgcagg gtcgatgcga cgcaatcgtc cgatccggag 540
ccgggactgt cgggcgtaca caaatcgccc gcagaagcgc ggccgtctgg accgatggct 600
gtgtagaagt actcgccgat agtggaacc gacgccccag cactcgctcc agggcaaagg 660
aatgagctcg ctgatcagcc tcgactgtgc cttctagtgt ccagccatct gttgtttgce 720
cctccccctg gccttccat tgtctgagta ggtgtcattc tattctgggg ggtggggtgg 780
ggcaggacag caaggggggag gattgggaag acaatagcag gcatgctggg gatgcggtgg 840
gctctatggc ttctgaggcg gaaagaacca gctgggggat atcaagctta tcgataccgt 900
cgacctcgag ggggggcccg gtaccagct tttgttccct ttagtgaggg ttaattgcgc 960
gcttggcgta atcatggtca tagctgtttc ctgtgtgaaa ttgttatccg ctcacaattc 1020
cacacaacat acgagccgga agcataaagt gtaaagcctg ggggtgcctaa tgagtgaagt 1080
aactcacatt aattgcgttg cgctcactgc ccgctttcca gtcgggaaac ctgtcgtgcc 1140
agctgcatta atgaatcggc caacgcgcgg ggagaggcgg tttgcgtatt gggcgctctt 1200
ccgcttcttc gctcactgac tcgctgcgct cggctgttcg gctgcggcga gcggtatcag 1260
ctcactcaaa ggccgtaata cggttatcca cagaatcagg ggataacgca ggaaagaaca 1320
tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa ggccgcgttg ctggcgtttt 1380
tccataggct ccgccccctt gacgagcatc acaaaaatcg acgctcaagt cagagggtgc 1440
gaaaccgcag aggactataa agataccagg cgtttccccc tggaagctcc ctcgctgcgt 1500
ctcctgttcc gacctgccc cttaccggat acctgtccgc ctttctccct tcgggaagcg 1560
tggcgctttc tcatagctca cgctgtaggt atctcagttc ggtgtaggtc gttegtccca 1620
agctgggctg tgtgcacgaa cccccgctt agcccgaccg ctgctgctta tccggttaact 1680
atcgtcttga gtccaacccg gtaagacacg acttatcgcc actggcagca gccactggta 1740
acaggattag cagagcgagg tatgtaggag gtgctacaga gttcttgaag tgggtggccta 1800
actacggcta actagaagg acagtatttg gtagctgcgc tctgctgaag ccagttacct 1860
tcggaaaaag agttggtagc tcttgatccg gcaaacaaac caccgctggt agcgggtggt 1920
tttttgtttg caagcagcag attacgcgca gaaaaaaagg atctcaagaa gatcctttga 1980
tcttttctac ggggtctgac gctcagtggg acgaaaactc acgttaaggg attttgggtca 2040
tgagattatc aaaaaggatc ttcacctaga tccttttaaa ttaaaaatga agtttttaaat 2100
caatctaaag tatatatgag taaacttggg ctgacagtta ccaatgctta atcagtgagg 2160

```

cacctatctc	agcgatctgt	ctatttcggt	catccatagt	tgcctgactc	cccgtcgtgt	2220
agataactac	gatacgggag	ggcttaccat	ctggccccag	tgctgcaatg	ataccgcgag	2280
acccacgctc	accggctcca	gatttatcag	caataaacca	gccagccgga	agggccgagc	2340
gcagaagtgg	tcctgcaact	ttatccgcct	ccatccagtc	tattaattgt	tgccgggaaag	2400
ctagagtaag	tagttcgcca	gttaatagtt	tgcgcaacgt	tggtgccatt	gctacaggca	2460
tcgtgggtgc	acgctcgctc	tttggtagtg	cttcattcag	ctccggttcc	caacgatcaa	2520
ggcgagttac	atgatcccc	atgttgtgca	aaaaagcggg	tagctccttc	ggtcctccga	2580
tcgttgtcag	aagtaagttg	gccgcagtg	tatcactcat	ggttatggca	gcactgcata	2640
attctcttac	tgtcatgcca	tccgtaagat	gcttttctgt	gactgggtgag	tactcaacca	2700
agtcattctg	agaatagtg	atgcggcgac	cgagttgctc	ttgcccggcg	tcaatacggg	2760
ataataccgc	gccacatagc	agaacttta	aagtgtcat	cattggaaaa	cgttcttcgg	2820
ggcgaaaact	ctcaaggatc	ttaccgctgt	tgagatccag	ttcgatgtaa	cccactcgtg	2880
cacccaactg	atcttcagca	tcttttactt	tcaccagcgt	ttctgggtga	gcaaaaacag	2940
gaaggcaaaa	tgccgcaaaa	aagggaataa	gggcgacacg	gaaatgttga	atactcatac	3000
tcttcctttt	tcaatattat	tgaagcattt	atcagggtta	ttgtctcatg	agcggataca	3060
tatttgaatg	tatttagaaa	aataaaca	taggggttcc	gcgcacattt	ccccgaaaag	3120
tgccacctaa	attgtaacg	ttaatatttt	gttaaaattc	gcgttaaatt	tttgttaa	3180
cagctcattt	tttaaccaat	aggccgaaat	cgcaaaatc	ccttataaat	caaaagaata	3240
gaccgagata	gggttgagtg	ttgttccagt	ttggaacaag	agtccactat	taaagaacgt	3300
ggactccaac	gtcaaagggc	gaaaaaccgt	ctatcagggc	gatggcccac	tacgtgaacc	3360
atcaccctaa	tcaagttttt	tggggtcgag	gtgccgtaaa	gcactaaatc	ggaaccctaa	3420
agggagcccc	cgatttagag	cttgacgggg	aaagccggcg	aacgtggcga	gaaaggaagg	3480
gaagaaaagc	aaaggagcgg	gcgctagggc	gctggcaagt	gtagcggcca	cgctgcgcgt	3540
aaccaccaca	ccgcgcgcgc	ttaatgcgcc	gctacagggc	gcgtcccat	cgccattcag	3600
gctgcgcaac	tggtgggaag	ggcgatcggt	gcgggcctct	tcgctattac	gccagctggc	3660
gaaaggggga	tgtgctgcaa	ggcgattaa	ttgggtaacg	ccagggtttt	cccagtcacg	3720
acgttgtaaa	acgacggcca	gtgagcgcg	gtaatacgac	tcactatagg	gcgaattgga	3780
gctccaccgc	ggtggcgggc	gcttgccaag	ggcctgtga	agcagggtac	cccaaaactc	3840
caacttgctg	tgccccag	ccccctagaa	ctgaagtgtt	tccaggactg	aacctctcac	3900
cccagaggcc	tcagaggtga	gcaggcctag	caggggccag	gcagtctgc	ggccactgac	3960
gtcggcctgg	gagggtaggac	aggcccccg	ctcagagcca	tctgtggttc	tggttagac	4020
ccatccctcc	tcgtctctgg	gcctcagggtg	cctcctctgc	agggtagggt	tgataaccac	4080
cgcgtgcct	ccctccccac	acagaacaag	gaaaacggcc	gggtgggctg	tcgtgggctg	4140
aattcaggga	ggatctcggt	gatgaggatg	gctcagtcga	gaagagccaa	gcagactgtg	4200
tccttggaat	agggtagcgg	gcagagactg	gccccaccac	acccacaccg	cactgcgggg	4260
ccctgctgtg	caacgggtcc	acacatcatc	agccagggtgc	ccggcctgcc	tttgccttc	4320
aacctcgctg	ccacagccac	tgagcttata	cgagattttg	gattttgtgc	ttttgtcat	4380
ttttatgaat	caacatctgt	caagcatcct	ggccttgaat	gtttaggcgt	gtgtgtgtgt	4440
gtgtgtgtgt	gtgtgtctca	ctctgtcgcc	cagactagag	tgcaatgggg	caatctctgt	4500
tcactgcaac	ctccacctcc	caggttcaag	cgattctcct	ccctcagcct	cccaagtagc	4560
tggtgattaca	ggcgcccgcc	accatgcgtg	gctaattgtt	gtatttttag	tagagcaagg	4620
gaacgttttag	gctttaagtg	atttgagct	ggttgggtctt	ctaggcctaa	acacctagag	4680
atcagctggc	cgctggcagc	agcatgtgta	gttgacaca	tgagctgatg	ctattttcca	4740
ctcccatcct	gccccatccc	catgcttccc	ctccctctgg	ctgggcaccc	ccactccctg	4800
gcacccaggg	gtggcccata	catccctccc	atagcctcta	ttcccatcgg	gtctccacct	4860
aagggttttt	tcctctgtct	acccttgcac	aagtctatc	accactcaga	tgctacttcc	4920
cctcaggatc	cagcactgat	gccgcctcct	ccatgaagcc	cccagatccc	tgccctgctc	4980
atgctccccg	ctgtcctcca	cacagtggcc	cagcccaactg	accgcctgtc	ccacacagcc	5040
taccatgtgg	tcagcaatt	gctatctgtc	actgatttat	ctctcccgca	ccctgtgacc	5100
tcctcctggg	caggacctt	tcctgggtcca	tctccttatg	aagaagctgt	actggcctgg	5160
gcgcgtggct	cacgcctgta	atcccagcac	tttgggaggc	caaggcaggt	ggatcacctg	5220
aggtcaggag	ttcgagacca	gcctggccaa	catggtgaaa	tcccatctct	actaaaaaaa	5280
atacaaaaat	tagcggggcg	tggtgtcggg	cgctgtaat	cccagctact	cgggaggctg	5340
aggcaggaga	atcgcttgaa	cccaggatgt	agaggttaca	gtgagctgag	actgtgccac	5400
tgactccag	cctgggcaac	agagtgaac	tctgtctcaa	aaaaaaaaaa	aaaaaaaaaa	5460
aaagaagcag	cagcagcagc	agcagcagca	gctgcacaca	tggggaaata	agagcaatca	5520
gacccttcta	attgccattt	ccataggaag	ttaggggagat	tactgttcca	gcagtggtca	5580
cattccggaa	aacagccgca	gaagaaatga	gaggcttctc	aagtcctgca	ttttcttttt	5640

tttttttttt	gaaatatata	atcaggcaat	ctcccaagcc	agaatggctc	agagagactc	5700
ccaagtccta	catttttcaag	ttaaaattca	ggcagtgaaa	ctccagagca	cacaactggg	5760
ccatggctgt	aaaccagcta	cgttctgaat	ctcagatcca	tgaaaacaca	gaggcagcct	5820
caccactccg	caggctgtga	acaggacgga	cggctggcta	ctggccatct	ccctgcctgc	5880
ctgttgatcc	actttcaaca	ggcctgagtc	tttctctgac	tcacccatca	ggtggacccc	5940
ctgcagcagg	agtgttcccc	gccacagaaa	gccgattccc	tcattctcatc	actcaatacc	6000
tatctactga	gtcctcctcc	tcagagtcca	gccctgcgct	gggggacacag	agaggagtca	6060
gctccccgct	gagacagaaa	ccaggtactg	gctggggcgg	gggactgcag	gagggggcag	6120
ggtgctgaag	aggggtccat	ggcagttggt	gatggatgag	ctgagcctgg	gaaggggaaca	6180
gaggtgggct	aggggtctgga	gacagacagg	gaagcctaga	gcagctgaag	aacgggaagt	6240
cagtgggaca	ggaagacagg	aggccgggtt	agaaccagat	ggccgtcagc	cttaactcca	6300
aggaattgaa	agccaacatg	gttgcatgga	tgatgctttg	gaagcctggg	gggcccgggtg	6360
ctagtcttcc	ctttctgaag	ataacacccc	gtcaaggacc	gttcacagtg	cggacagccc	6420
tggcagcact	taaaaggcac	cagaagcgaa	tgagaagcag	aatcgttgat	gcgtctgagc	6480
acgtgcacaa	agcagacacg	gttctcagga	ctcgcattcc	ttcaccact	tcattcccac	6540
aacatcccta	agagggaat	acttcaccga	tgaggagacc	gaggcacaga	gcagttatgg	6600
cacgtccccg	aggtcacaca	gtcagagatg	gaaccggac	ttgaccccag	gcaggctggc	6660
tccagagcat	ctccagtgtc	ccaccacgat	gctggagggtc	tggggctctt	tctcagcagt	6720
tcaactcaga	cccagactcc	acagcagggc	ctggggggcca	atcagaatag	gccctctcac	6780
ccaaggggct	aggcataaga	cctgaacctg	aagcacctct	cgctcaaaca	agcagctcag	6840
gaacccccaa	atgcagtgtg	aataagaccc	aggtcatacc	catgcctggc	tgtcccaggg	6900
atgtctatgg	tctccaagaa	taagctaaga	ataagaatag	gggctgctcg	ttaggcagcc	6960
cagcaagaga	gcacctgtta	gccggagcct	cccatgaacc	tttcggcaat	gttctttgtc	7020
agcggcggtg	ccttttcttc	tgacgggtcc	cagtcacaacc	aaacccagcc	aacaggaagc	7080
agcctggatt	gtgctatcat	tcagtcagtc	gacaaacgct	gaatctatac	tgcacactct	7140
gcacctgatt	ccctactgga	aaccaagctt	cttccaggga	acgggggtggg	acaggaactg	7200
aggggtcatg	tgccctgggc	tgggccgggg	aacagggtca	ggacagccct	ccactgccag	7260
gcctggagtg	taccctgcac	ctaagcccgt	ggtgatgcac	ctgtcaccgc	aaaacttct	7320
acagtgatca	ggggaacttc	taccgtgatt	taggcaaacg	agtgttggct	gtgtgatgga	7380
gatggtgatt	ctggtttaca	caagactcac	aaactcatca	atctggacta	aaggacaatc	7440
ctaaaagaat	cgattctgcc	cacagagacg	tggtatgatc	tgagtctttc	tcaactccaa	7500
agccagaaca	gatactgcct	accctctggg	ggaaggatcc	cttctctccc	caatgcaagt	7560
gagacaacat	gggggtgtggc	tctctagaac	tagtggatcc	cccgggaccg	gtccccccac	7620
tccactgcc	gtaagcagtg	ggttctctag	ttagccagag	agctctgctt	atatagacct	7680
cccaccgtac	acgcctaccg	cccatttgcg	tcaattggggc	ggagttgtta	cgacattttg	7740
gaaagtcccc	ttgatttttg	tgccaaaaca	aactcccatt	gacgtcaatg	gggtggagac	7800
ttggaaatcc	ccgtgagtca	aaccgctatc	cacgcccatt	gatgtactgc	caaaaccgca	7860
tcaccatggt	aatagcgatg	actaatacgt	agatgtactg	ccaagtagga	aagtcccata	7920
aggtcatgta	ctgggcataa	tgccaggcgg	gccatttacc	gtcattgacg	tcaatagggg	7980
gcgtacttgg	catatgatac	acttgatgta	ctgccaaagt	ggcagtttac	cgtaaatagt	8040
ccaccatttg	acgtcaatgg	aaagtcccta	ttggcggttac	tatgggaaca	tacgtcatta	8100
ttgacgtcaa	tgggcggggg	tcgttgggcg	gtcagccagg	cgggccattt	accgtaagtt	8160
atgtaacgcg	gaactccata	tatgggctat	gaactaatga	ccccgtaatt	gattactatt	8220
aataactagt	caataatcaa	tgtcaacgcc	cgggctgcag	gaattctacc	gggtagggga	8280
ggcgcttttc	ccaaggcagt	ctggagcatg	cgttttagca	gccccgctgg	cacttggcgc	8340
atcacaagtg	gctcttgccc	tcgcacacat	tccacatcca	ccggtagcgc	caaccggctc	8400
ccttcttttg	tggcccttcc	gcgccacctt	ctactcctcc	cctagtcagg	aagttcccc	8460
ccgccccgca	gctcgcgtcg	tgcaaggacgt	gacaaatgga	agtagcacgt	ctcactagtc	8520
tcgtgcagat	ggacaagcac	cgctgagcaa	tggaagcggg	taggcctttg	gggcagcggc	8580
caatagcagc	ttggctcctt	cgctttcttg	gctcagaggc	tgggaagggg	tgggtccggg	8640
ggcgggctca	ggggcgggct	caggggcggg	gcgggcgcga	aggtcctccg	gaccgggcat	8700
tctgcacgct	tcaaaaagcg	acgtctgccc	cgtgtgtctc	ctcttctctc	tctccgggccc	8760
ttcgactctg	atgaaaaagc	ctgaactcac	cgcgacgtct	gtcgaagaagt	ttctgatcga	8820
aaagttcgac	agcgtctccg	acctgatgca	gctctcggag	ggcgaagaat	ctcgtgcttt	8880
cagcttcgat	gtaggagggc	gtggatatgt	cctgcgggta	aatagctgcg	ccgatggttt	8940
ctacaaagat	cgttatgttt	atcggcactt	tgcacggccc	gcgctcccga	ttccggaagt	9000

```

gcttgacatt ggggaattca gcgagagcct gacctattgc atctcccgcc gtgcacaggg 9060
tgtcacgttg caagacctgc ctgaaaccga actgcccgtt gttctgcagc cggctcgcgga 9120
ggccatggat g                                     9131

```

<210> 84  
 <211> 1121  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: Synthetic  
 EYFP sequence

```

<400> 84
cttgtacagc tcgtccatgc cgagagtgat cccggcggcg gtcacgaact ccagcaggac 60
catgtgatcg cgcttctcgt tggggtcttt gctcagggcg gactggtagc tcaggtagtg 120
gttgctcgggc agcagcacgg ggccgtcgcc gatgggggtg ttctgctggt agtggtcggc 180
gagctgcacg ctgccgtcct cgatgttggt gcggatcttg aagttcacct tgatgccgtt 240
cttctgcttg tcggccatga tatagacgtt gtggctgttg tagttgtact ccagcttgtg 300
ccccaggatg ttgccgtcct ccttgaagtc gatgcccttc agctcgatgc gggtcaccag 360
ggtgtcgccc tcgaacttca cctcggcgcg ggtctttag ttgccgtcgt ccttgaagaa 420
gatggtgcgc tcctggacgt agccttcggg catggcggac ttgaagaagt cgtgctgctt 480
catgtggtcg gggtagcggg cgaagcactg caggccgtag ccgaagggtg tcacgagggg 540
gggccagggc acgggcagct tgccggtggt gcagatgaac ttcagggtca gcttgccgta 600
ggtggcatcg ccctcgccct cgccggacac gctgaacttg tggccgttta cgtcgccgtc 660
cagctcgacc aggatgggca ccaccccggt gaacagctcc tcgcccttgc tcaccatggt 720
ggcgaccggt ggatccttct agagtcgacg atcgaagtta gcttgatttg acagtggctg 780
ggggttcgcg cgccgggttt tataggaagc cacagcggcc actcgagcca taaaaggcaa 840
cttttagaac ggcggggggg gattggattc gagtcgttta ttcaccggcc ttgccgcaca 900
gtgcagcatt tttttacccc ctctcccctc cttttgcggg ggaaaaaaaa aaaaaaaaaa 960
aaaaaaaaag aagagaaaaa aagcgagcga gagagaaagc gagattgagg aagaggatga 1020
agagtttggc gatgggtgct ggttccgtag gccagatgg acaagaatag cccccgccct 1080
tgcgacagt atccattca gtgactcaga tcagatcaag c                                     1121

```

<210> 85  
 <211> 10557  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: Synthetic  
 pKI-CMV-SD-Vanilloid-YFP sequence

```

<400> 85
tttttaccct ctctcccctc cttttgcggg ggaaaaaaaa aaaaaaaaaa aaaaaaggag 60
aagagaaaaa aagcgagcga gagagaaagc gagattgagg aagaggatga agagtttggc 120
gatgggtgct ggttccgtag gccagatgg acaagaatag cccccgccct tgcggacagt 180
atccattca gtgactcaga tcagatcaag ccggccgctt gccaagggcc ctgtgaagca 240
gggtacccca aaactccaac ttgtgtctgc cccagcccc ctagaactga agtgtttcca 300
ggactgaacc tctcaccaca gaggcctcag aggtgagcag gcctagcagg gccaggcag 360
tcctgcggcc actgacgtcg gcctgggagg gtggacaggc ccccggtca gagccatctg 420
tggttctggg ttagaccat ccctcctcgt ctctgggct cagggtgcctc ctctgcaggg 480
tgaggttgat aaccaccgcg ctgcctccct cccacacag aacaaggaaa acggccgggg 540
tggtgtcgt gggctgaatt caggaggat ctcggtgatg aggatggctc agtccagaag 600
agccaagcag actgtgtccc tggaataggg tgagcggcag agactggccc acccacacc 660
acaccgcact gcggggccct gctgtgcaac ggtccacac atcatcagcc aggtgcccgg 720

```



cctgcctttg	ctcttcaacc	tcgctgccac	agccactgag	cttatacgag	atthttggatt	780
ttgtgctttt	tgtcattttt	atgaatcaac	atctgtcaag	catcctggcc	ttgaatgttt	840
agggcgtgtg	gtgtgtgtgt	gtgtgtgtgt	gtctcactct	gtcgcccaga	ctagagtga	900
atggggcaat	ctctgttcac	tgcaacctcc	acctcccagg	ttcaagcgat	tctcctccct	960
cagcctccca	agtagctggg	attacaggcg	cccgccacca	tgcgtggcta	atgtttgtat	1020
ttttagtaga	gacaggggaa	gtttaggctt	taagtgattt	ggagctgggt	ggctcttctag	1080
gcctaaacac	ctagagatca	gctggccgct	ggcagcagca	tgtgtagttg	gacacatgag	1140
ctgatgctat	tttccactcc	catcctgccc	catccccatg	cttccccctc	ctctggctgg	1200
gcacccccac	tccctggcac	ccaggggtgg	cccatcacatc	cctccccatag	cctctattcc	1260
catcggtct	ccacctaagg	gctttttcct	ctgtctaccc	ttgcacaagt	cctatcacca	1320
ctcagatgct	acttccccctc	aggatccagc	actgatgccg	cctcctccat	gaagccccc	1380
gatccctgcc	ctgctcatgc	tccccgctgt	cctccacaca	gtggcccagc	ccactgaccg	1440
cctgtcccac	acagcctacc	atgtggtcca	gcaattgcta	tctgtcactg	atthtatctct	1500
cccgcaccct	gtgacctcct	cctgggcagg	gacctttcct	gggtccatctc	cttatgaaga	1560
agctgtactg	gcctggggcg	gtggctcacg	cctgtaatcc	cagcactttg	ggaggccaag	1620
gcacagggat	cacctgaggt	caggagttcg	agaccagcct	ggccaacatg	gtgaaatccc	1680
atctctacta	aaaaaaatac	aaaaattagc	cgggcgtggg	gtcggggcgcc	tgtaatccca	1740
gctactcggg	aggctgaggg	aggagaatcg	cttgaaccca	ggatgtagag	gttacagtga	1800
gctgagactg	tgccactgca	ctccagcctg	ggcaacagag	tgaaactctg	tctcaaaaa	1860
aaaaaaaaaa	aaaaaaaaaag	aagcagcagc	agcagcagca	gcagcagctg	cacacatggg	1920
gaaataagag	caatcagacc	cttctaattg	ccattttccat	aggaagttag	ggagattact	1980
gttcacgacg	tgggtacatt	ccggaaaaca	gcccagagaag	aaatgagagg	cttctcaagt	2040
cctgcatttt	cttttttttt	tttttttgaaa	tatataatca	ggcaatctcc	caagccagaa	2100
tggtctcagag	agactcccaa	gtcctacatt	ttcaagttaa	aattcaggca	gtgaaactcc	2160
agagcacaca	actggggccat	ggctgtaaac	cagctacgtt	ctgaatctca	gatccatgaa	2220
aacacagagg	cagcctcacc	actccgcagg	ctgtgaacag	gacggacggc	tggtctactgg	2280
ccatctccct	gcctgcctgt	tgatccactt	tcaacaggcc	tgagtctttc	tctgactcac	2340
ccatcagggt	gacccctgc	agcaggagtg	ttcccggcca	cagaaagccg	attcctcat	2400
ctcatcactc	aatacctatc	tactgagttc	ttcctcctcag	agtccagccc	tgcgctgggg	2460
gcacagagag	gagtcagctc	cccgtcgaga	cagaaaccag	gtactggctg	ggcggaggga	2520
ctgcaggagg	gggcagggtg	ctgaagaggg	gtccatggca	gttgggtgat	gatgagctga	2580
gcctgggaag	ggaacagagg	tgggctaggg	tctggagaca	gacaggggag	cctagagcag	2640
ctgaagaacg	ggaagtcagt	gggacaggaa	gacaggaggc	cgggttagaa	ccagatggcc	2700
gtcagcctta	actccaagga	attgaaagcc	aacatggttg	catggatgat	gctttggaag	2760
cctggggggc	cgggtgctag	tcttcccttt	ctgaagataa	caccccgta	aggaccgttc	2820
acagtgcgga	cagccctggc	agcacttaa	aggcaccaga	agcgaatgag	aagcagaatc	2880
gttgatgcgt	ctgagcacgt	gcacaaagca	gacacggttc	tcaggactcg	cattccttca	2940
cccacttcat	tcccacaaca	tcccctaagag	ggaaatactt	caccgatgag	gagaccgagg	3000
cacagagcag	ttatggcacg	tccccgaggt	cacacagtca	gagatggaac	ccggacttga	3060
cccaggcagc	gctggctcca	gagcatctcc	agtgtctccac	cacgatgctg	gaggtctggg	3120
gctctttctc	agcagttcaa	ctcagaccaca	gactcccacag	cagggcctgg	gggccaatca	3180
gaataggccc	tctcacccaa	ggggctaggg	ataagacctg	aacctgaagc	acctctcgct	3240
caaacaagca	gctcagggaa	cccaaaatgc	agtgtgaata	agaccaggt	catacccatg	3300
cctggctgtc	ccagggatgt	ctatggtctc	caagaataag	ctaagaataa	gaataggggc	3360
tgctcgttag	gcagcccagc	aagagagcac	ctgttagccg	gagcctccca	tgaacctttc	3420
ggcaatgttc	tttgtcagcg	gcgttgccct	ttcctctgac	gggtcccagt	ccaaccaaac	3480
ccagccaaca	ggaagcagcc	tggattgtgc	tatcattcag	tcagtcgaca	aacgctgaat	3540
ctatactgca	cactctgcac	ctgattccct	actggaaacc	aagcttcttc	cagggaacgg	3600
ggtgggacag	gaactgaggg	gtcatgtgcc	ctgggctggg	ccggggaaca	gggtcaggac	3660
agccctccac	tgccaggcct	ggagtgtacc	ctgcacctaa	gcccgtgggt	atgcacctgt	3720
caccgcaaaa	cattctacag	tgatcagggg	aacttctacc	gtgatttagg	caaacgagtg	3780
ttggctgtgt	gatggagatg	gtgattctgt	tttacacaag	actcacaac	tcatcaatct	3840
ggactaaaagg	acaatcctaa	agaatcgat	tctgccaca	gagacgtggg	atgatctgag	3900
tctttctcaa	ctccaaagcc	agaacagata	ctgcctaccc	tctgggggaa	ggatcccttc	3960
tctccccaat	gcaagtgaga	caacatgggg	tgtggctctc	tagaactagt	ggatcccccg	4020
ggaccgggtc	cccactcac	ctgccagtaa	gcagtgggtt	ctctagttag	ccagagagct	4080
ctgcttatat	agacctccca	ccgtacacgc	ctaccgcccc	tttgcgtcaa	tggggcggag	4140
ttgttacgac	atthttggaaa	gtcccgttga	ttttggtgcc	aaaacaaact	cccattgacg	4200

tcaatgggggt	ggagacttgg	aaatccccgt	gagtcaaacc	gctatccacg	cccattgatg	4260
tactgccaaa	accgcatcac	catggtaata	gcgatgacta	atacgtagat	gtactgcaa	4320
gtaggaaagt	cccataaggt	catgtactgg	gcataatgcc	aggcggggcca	tttaccgtca	4380
ttgacgtcaa	tagggggcgt	acttggcata	tgatacactt	gatgtactgc	caagtgggca	4440
gtttaccgta	aatagtccac	ccattgacgt	caatggaaag	tccctatttg	cgttactatg	4500
ggaacatacg	tcattattga	cgtcaatggg	cgggggtcgt	tgggcggtca	gccaggcggg	4560
ccatttaccg	taagttatgt	aacgcggaac	tccatatatg	ggctatgaac	taatgacccc	4620
gtaattgatt	actattaata	actagtcaat	aatcaatgtc	aacgcccggg	ctgcaggaat	4680
tctaccgggt	aggggaggcg	cttttcccaa	ggcagtcctg	agcatgcgct	ttagcagccc	4740
cgctggcact	tggcgcatca	caagtggcct	ctggcctcgc	acacattcca	catccaccgg	4800
tagcgccaac	cggctccctt	ctttggtggc	cccttcgcgc	caccttctac	tcctccccta	4860
gtcaggaagt	cccccccg	cccgcagctc	gcgtcgtgca	ggacgtgaca	aatggaagta	4920
gcacgtctca	ctagtctcgt	gcagatggac	aagcaccgct	gagcaatgga	agcgggtagg	4980
cctttggggc	agcggccaat	agcagcttgg	ctccttcgct	ttctgggctc	agaggctggg	5040
aaggggtggg	tccggggg	ggctcagggg	cgggctcagg	ggcggggcgg	gcgcgaaggt	5100
cctccggacc	cggcattctg	cacgcttcaa	aagcgacgt	ctgccgcgct	gttctcctct	5160
tcctcatctc	cgggccttcg	acctgcatga	aaaagcctga	actcaccgcg	acgtctgtcg	5220
agaagtttct	gatcgaaaag	ttcgacagcg	tctccgacct	gatgcagctc	tcggagggcg	5280
agaatctcgt	tgctttcagc	ttcgatgtag	gagggcgctg	atatgtcctg	cgggtaaaata	5340
gctgcgccga	tggtttctac	aaagatcggt	atgtttatcg	gcactttgca	tcggccgcgc	5400
tcctcgattcc	ggaagtgtct	gacattgggg	aattcagcga	gagcctgacc	tattgcatct	5460
cccgcgctgc	acagggtgtc	acgttgcaag	acctgcctga	aaccgaactg	cccgtgttct	5520
tgacgccggt	cgcgaggcc	atggatgcga	tcgctgcggc	cgatcttagc	cagacgagcg	5580
ggttcggccc	attcggaccg	caaggaatcg	gtcaatacac	tacatggcgt	gatttcatat	5640
gcgcgattgc	tgatcccat	gtgtatcact	ggcaaactgt	gatggacgac	accgtcagtg	5700
cgctcgtcgc	gcaggctctc	gatgagctga	tgttttgggc	cgaggactgc	cccgaagtcc	5760
ggcacctcgt	gcacgcggat	ttcggtccca	acaatgtcct	gacggacaat	ggccgcataa	5820
cagccggtcat	tgactggagc	gaggcgatgt	tcggggattc	ccaatacgag	gtcgccaaca	5880
tcctcttctg	gaggcgtgg	ttggcttgta	tgggacgaca	gacgcgctac	ttcgagcgga	5940
ggcatccgga	gcttgacgga	tcgccgcggc	tcggggcgta	tatgtctccg	attggctctg	6000
accaactcta	tcagagcttg	gttgacggca	atttcgatga	tcagcttgg	gcgcagggct	6060
gatgcgacgc	aatcgtccga	tcgggagccg	ggactgtcgg	gcgtacacaa	atcgcccgcga	6120
gaagcgcggc	cgtctggacc	gatggctgtg	tagaagtact	cgccgatagt	ggaaaccgac	6180
gccccagcac	tcgtccgagg	gcaaagggaat	gagctcgtcg	atcagcctcg	actgtgcctt	6240
ctagttgccca	gccattctgt	gtttgcccc	cccccgctgc	ttcccatgtg	ctgagtaggt	6300
gtcattctat	tctggggggg	gggggtgggg	aggacagcaa	gggggaggat	tgggaagaca	6360
atagcaggca	tgctggggat	gcgggtgggct	ctatggcttc	tgaggcggaa	agaaccagct	6420
gggggataatc	aagcttatcg	ataccgtcga	cctcgagggg	gggcccggta	cccagctttt	6480
gttcccttta	gtgaggggta	attgcgcgct	tggcgtaatc	atgggtcatag	ctgtttcctg	6540
tgtgaaattg	ttatccgctc	acaattccac	acaacatacg	agccggaagc	ataaagtgtg	6600
aagcctgggg	gtgcctaatga	gtgagctaag	tcacattaat	tgctttgcgc	tcactgccc	6660
ctttccagtc	gggaaacctg	tcgtgccagc	tgcattaatg	aatcggccaa	cgcgcgggga	6720
gaggcggttt	gcgtattggg	cgctcttcog	cttctcgtct	cactgactcg	ctgcgctcgg	6780
tcgttcgggt	gcggcgagcg	gtatcagctc	actcaaaggc	ggtaatacgg	ttatccacag	6840
aatcagggga	taacgcagga	aagaacatgt	gagcaaaaag	ccagcaaaaag	gccaggaacc	6900
gtaaaaaggc	cgcgtttgctg	gcgtttttcc	ataggctccg	ccccctgac	gagcatcaca	6960
aaaatcgacg	ctcaagtacg	aggtggcgaa	acccgacagg	actataaaga	taccaggcgt	7020
ttccccctgg	aagctccctc	gtgcgctctc	ctgttcgcgac	cctgccgctt	accggatacc	7080
tgtccgcctt	tctcccttcg	ggaagcggtg	cgctttctca	tagctcacgc	tgtaggtatc	7140
tcagttcggt	gtaggtcgtt	cgctccaagc	tgggctgtgt	gcacgaaccc	cccgttcagc	7200
ccgaccgctg	cgccttatcc	ggtaactatc	gtcttgagtc	caaccgggta	agacacgact	7260
tatcgccact	ggcagcagcc	actggtaaca	ggattagcag	agcgaggtat	gtaggcgggtg	7320
ctacagagtt	cttgaagtgg	tggcctaact	acggctacac	tagaaggaca	gtatttggtg	7380
tctgcgctct	gctgaagcca	gttaccttcg	gaaaaagagt	tggtagctct	tgatccggca	7440
aacaaaccac	cgctggtagc	ggtaggtttt	ttgtttgcaa	gcagcagatt	acgcgcagaa	7500
aaaaaggatc	tcaagaagat	cctttgatct	tttctacggg	gtctgacgct	cagtggaaacg	7560
aaaactcacg	ttaagggatt	ttggtcatga	gattatcaaa	aaggatcttc	acctagatcc	7620
ttttaaatga	aaaatgaagt	tttaaatcaa	tctaaagtat	atatgagtaa	acttggtctg	7680

```

acagttacca atgcttaatc agtgaggcac ctatctcagc gatctgtcta tttcgttcat 7740
ccatagttgc ctgactcccc gtcgtgtaga taactacgat acgggagggc ttaccatctg 7800
gccccagtgc tgcaatgata ccgcgagacc cagctcacc ggctccagat ttatcagcaa 7860
taaaccagcc agccggaagg gccgagcgca gaagtgggcc tgcaacttta tccgcctcca 7920
tccagtctat taattgttgc cgggaagcta gagtaagtag ttcgccagtt aatagtttgc 7980
gcaacgttgt tgccattgct acaggcatcg tgggtgtcacg ctcgctcgttt ggtatggctt 8040
cattcagctc cggttcccaa cgatcaaggc gagttacatg atccccatg ttgtgcaaaa 8100
aagcggttag ctccctcggg cctccgatcg ttgtcagaag taagttggcc gcagtgttat 8160
cactcatggg tatggcagca ctgcataatt ctcttactgt catgccatcc gtaagatgct 8220
tttctgtgac tgggtgagta tcaaccaagt cattctgaga atagtgtatg cggcgaccga 8280
gttgctcttg cccggcgta atacgggata ataccgcgcc acatagcaga actttaaaag 8340
tgctcatcat tggaaaacgt tcttcggggc gaaaactctc aaggatctta ccgctgttga 8400
gatccagttc gatgtaaccc actcgtgcac ccaactgatc ttcagcatct tttactttca 8460
ccagcgtttc tgggtgagca aaaacaggaa ggcaaaatgc cgcaaaaaag ggaataaggg 8520
cgacacggaa atgttgaaata ctcatactct tcccttttca atattattga agcatttatc 8580
agggttattg tctcatgagc ggatacatat ttgaatgtat ttagaaaaat aaacaaatag 8640
gggttccgag cacatttccc cgaaaagtgc cacctaaatt gtaagcgtaa atattttgtt 8700
aaaattcgcg ttaaatTTTT gttaaatcag ctcatTTTTt aaccaatagg ccgaaatcgg 8760
caaaatccct tataaatcaa aagaatagac cgagataggg ttgagtgttg ttccagtttg 8820
gaacaagagt ccactattaa agaacgtgga ctccaacgtc aaagggcgaa aaaccgtcta 8880
tcagggcgat ggcccactac gtgaaccatc accctaataca agtttttttg ggtcgaggtg 8940
ccgtaaagca ctaaatcggg accctaagg gagccccga tttagagctt gacggggaaa 9000
gccggcgaaac gtggcgagaa aggaagggaa gaaagcgaaa ggagcgggcg ctagggcgct 9060
ggcaagtgtg gcggtcacgc tgcgcgtaac caccacaccc gccgcgctta atgcgcgct 9120
acagggcgcg tccattcgc cattcaggct gcgcaactgt tgggaagggc gatcggtgcg 9180
ggcctcttcg ctattacgcc agctggcgaa agggggatgt gctgcaaggc gattaagttg 9240
ggtaacgcca gggttttccc agtcacgacg ttgtaaaacg acggccagtg agcgcgcgta 9300
atacgactca ctatagggcg aattggagct ccaccgcgtt ggcgcccgca acagacatga 9360
taagatacat tagtagtatt ggacaaacca caactagaat gcagtgaaaa aaatgcttta 9420
tttgtgaaat ttgtgatgct attgctttat ttgtaacat tataagctgc aataaacaag 9480
ttgggggtgg cgaagaactc cagcatgaga tccccgcgt ggaggatcat ccagccggcg 9540
tcccggaaaa cgattccgaa gcccaacctt tcatagaagg cggcggtgga atcgaaatct 9600
cgtgatggca ggttgggcgt cgcttggctg gtcatttcga agctttactt gtacagctcg 9660
tccatgccga gagtgatccc ggcgggcggt acgaactcca gcaggacct gtgatcgcg 9720
ttctcgttgg ggtctttgct cagggcgagc tggtagctca ggtagtgggt gtcgggcagc 9780
agcacggggc cgtcgccgat ggggggtgtt tgctggtagt ggtcggcgag ctgcacgctg 9840
ccgtcctcga tgttgtggcg gatcttgaag ttaccttga tgcggttctt ctgcttgtcg 9900
gccatgatat agacgttgtg gctgtttag ttgtactcca gcttgtgccc caggatgttg 9960
ccgtcctcct tgaagtcgat gcccttcagc tcgatgcggt tcaccagggt gtcgccctcg 10020
aacttcacct cggcgcggtt cttgtagtgt ccgtcgtcct tgaagaagat ggtgcgctcc 10080
tggacgtagc cttcgggcat ggcggaactt aagaagtct gctgcttcat gtggtcgggg 10140
tagcgggagc agcactgcag gccgtagccg aaggtggtca cgagggtggg ccagggcacg 10200
ggcagcttgc cggtggtgca gatgaacttc agggtcagct tgccgtaggt ggcacgccc 10260
tcgccctcgc cggacacgct gaacttgttg ccgtttacgt cgccgtccag ctcgaccagg 10320
atgggcacca ccccggtgaa cagctcctcg cccttgetca ccattggtgg gaccggtgga 10380
tccttctaga gtcgacgatc gaagttagct tgatttgaca gtggctgggg gttgcgccg 10440
cgggttttat aggaagccac agcggccact cgagccataa aaggcaactt taggaacggc 10500
ggggggtgat tggattcgag tcgtttattc accggccttg ccgcacagtg cagcatt 10557

```

<210> 86

<211> 3799

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic nucleotide sequence for Vanilloid region of homology

&lt;400&gt; 86

ttcaagcgat	tctcctgcct	cagcctcccc	agtagctggg	attacaggcg	cccgcacca	60
cgcccggcta	atTTTTgtat	TTTTTTtagt	agagatggga	tttcaccatg	ttggccaggc	120
tggtctcgaa	ctcctgacct	caggtgatcc	acctgccttg	gcctcccaaa	gtgctgggat	180
tacaggcgtg	agccacgcgc	ccaggccagt	acagcttctt	cataaggaga	tggaaccagga	240
aagggtccctg	cccaggagga	ggtcacaggg	tgccgggagag	ataaatcagt	gacagatagc	300
aattgctgga	ccacatggta	ggctgtgtgg	gacaggcggg	cagtgggctg	ggccactgtg	360
tgaggagacag	cggggagcat	gagcagggca	gggatctggg	ggcttcatgg	aggaggcggc	420
atcagtgcctg	gatcctgagg	ggaagtagca	tctgagtggt	gataggactt	gtgcaagggt	480
agacagagga	aaaagccctt	aggtggagac	ccgatgggaa	tagaggctat	gggagggatg	540
tatggggccac	ccctgggtgc	cagggagtgg	ggatgccag	ccagagggag	gggaagcatg	600
gggatggggc	aggatgggag	tggaaaatag	catcagctca	tgtgtccaac	tacacatgct	660
gctgccagcg	gccagctgat	ctctaggtgt	ttaggcctag	aagaccaacc	agctccaaat	720
cacttaaagc	ctaaacgttc	cctgtctcta	ctaaaaatac	aaacattagc	cacgcattgt	780
ggcgggcgcc	tgtaatccca	gctacttggg	agggtagggg	aggagaatcg	cttgaacctg	840
ggaggtggag	gttgcatgta	acagagattg	ccccattgca	ctctagtctg	ggcgacagag	900
tgagacacac	acacacacac	acacacacac	acgcctaaac	attcaaggcc	aggatgcttg	960
acagatgttg	attcataaaa	atgacaaaaa	gcacaaaatc	caaaatctcg	tataagctca	1020
gtggctgtgg	cagcgagggt	gaagagcaaa	ggcaggccgg	gcacctggct	gatgatgtgt	1080
ggaccgcgtg	cacagcaggg	ccccgcagtg	cggtgtgggt	gtgggtgggc	cagtctctgc	1140
cgcctaccct	attccaggga	cacagtctgc	ttggctcttc	tggactgagc	catcctcacc	1200
accagatcc	tccctgaatt	cagcccacga	cagccacccc	ggcgttttcc	cttgttctgt	1260
gtggggaggg	aggcagcgcg	gtggttatca	acctcacctc	gcagaggagg	cacctgaggc	1320
ccagagacga	ggagggatgg	gtctaaccac	gaaccacaga	tggctctgag	ccgggggcct	1380
gtccaccctc	ccaggccgac	gtcagtggcc	gcaggactgc	ctgggcccct	ctaggcctgc	1440
tcacctctga	ggcctctggg	gtgagaggtt	cagtcctgga	aacacttcag	ttctaggggg	1500
ctggggggcag	cagcaagtgg	gagttttggg	gtaccctgct	tcacagggcc	cttgggcaag	1560
ggcgcctcag	atctagagag	ccacacccca	tggtgtctca	cttgcatagg	ggagagaagg	1620
gatccttccc	ccagagggtg	ggcagtatct	gttctggctt	tggagttgag	aaagactcag	1680
atcataccac	gtctctgtgg	gcagaatcga	ttcttttagg	attgtccttt	agtccagatt	1740
gatgagtttg	tgagtcttgt	gtaaaccaga	atcaccatct	ccatcacaca	gccaacactc	1800
gtttgcctaa	atcacggtag	aagttcccct	gatcactgta	gaatgttttg	cgggtgacagg	1860
tgcatcacca	cgggcttagg	tgcagggtac	actccaggcc	tggcagtggg	gggctgtcct	1920
gacctgttcc	cccggcccag	cccaggccac	atgacccctc	agttcctgtc	ccaccccggt	1980
ccctggaaga	agcttggttt	ccagtagggg	atcagggtga	gagtggtgag	tatagattca	2040
gcgtttgtcg	actgactgaa	tgatagcaca	atccaggctg	cttctgtgtg	gctgggtttg	2100
gttgagactgg	gacccgtcag	aggaaaaggc	aacgcgcgtg	acaaagaaca	ttgccgaaag	2160
gttcatggga	ggctccggct	aacagggtgt	ctcttgctgg	gctgcctaac	gagcagcccc	2220
tattcttatt	cttagcttat	tcttgagagc	catagacatc	cctgggacag	ccaggcatgg	2280
gtatgacctg	ggcttatttc	acactgcatt	ttggggttcc	tgagctgctt	gtttgagcga	2340
gaggtgcttc	aggttcaggt	cttatgccta	gccccttggg	tgagaggggc	tattctgatt	2400
ggcccccagg	ccctgctgtg	gagtcctggg	ctgagttgaa	ctgctgagaa	agagccccag	2460
acctccagca	tcgtgggtgga	gcactggaga	tgctctggag	ccagcctgcc	tggggtcaag	2520
tccgggttcc	atctctgact	gtgtgacctc	ggggacgtgc	cataactgct	ctgtgcctcg	2580
gtctcctcat	cggggaagta	tttccctctt	agggatgttg	tgggaatgaa	gtgggtgaag	2640
gaatgcgagt	cctgagaacc	gtgtctgctt	tgtgcacgtg	ctcagacgca	tcaacgattc	2700
tgcttctcat	tcgcttctgg	tgccctttta	gtgtgtccag	ggctgtccgc	actgtgaacg	2760
gtccttgacg	gggtgttatc	ttcagaaaagg	gaagactagc	acccggcccc	ccaggcttcc	2820
aaagcatcat	ccatgcaacc	atgttggtct	tcaattcctt	ggagttaagg	ctgacggcca	2880
tctggttcta	acccggcctc	ctgtcttcc	gtcccactga	cttcccgttc	ttcagctgct	2940
ctaggcttcc	ctgtctgtct	ccagacccta	gcccacctct	gttcccttcc	caggctcagc	3000
tcacccatca	ccaactgcca	tggacccctc	tcagcacc	tgccccctcc	tgcagctcct	3060
ccgcccagcc	agtacctggt	ttctgtctca	gcggggagct	gactcctctc	tgtgccccca	3120
gcgcagggtc	ggactctgag	gaggaggact	cagtagatag	gtattgagtg	atgagatgag	3180
ggaatcggct	ttctgtggcc	gggaacactc	ctgctgcagg	gggtccacct	gatgggtgag	3240
tcagagaaag	actcaggcct	gttgaaagtg	gatcaacagg	caggcaggga	gatggccagt	3300
agccagccgt	cgtcctgttt	cacagcctgc	ggagtgggtg	ggctgcctct	gtgttttcat	3360

102/102

ggatctgaga	ttcagaacgt	agctggttta	cagccatggc	ccagttgtgt	gctctggagt	3420
ttcactgcct	gaattttaac	ttgaaaatgt	aggacttggg	agtctctctg	agccattctg	3480
gcttgggaga	ttgcctgatt	atatatttca	aaaaaaaaaa	aaaagaaaat	gcaggacttg	3540
agaagcctct	catttcttct	gcggctgttt	tccggaatgt	agccactgct	ggaacagtaa	3600
tctccctaac	ttcctatgga	aatggcaatt	agaagggctc	gattgctctt	atttcccat	3660
gtgtgcagct	gctgctgctg	ctgctgctgc	tgettctttt	tttttttttt	tttttttttt	3720
tgagacagag	tttactctg	ttgcccaggc	tggagtgcag	tggcacagtc	tcagctcact	3780
gtaacctcta	catcctggg					3799



Creation date: 09-08-2004  
Indexing Officer: SGIRON - SERGIO GIRON  
Team: OIPEBackFileIndexing  
Dossier: 10708420

Legal Date: 07-15-2004

No.	Doccode	Number of pages
1	PEFR	5
2	OATH	2
3	A.PE	1
4	CLM	5
5	REM	1
6	IDS	4
7	FOR	68
8	FOR	11
9	FOR	22
10	FOR	16
11	NPL	12
12	NPL	12

Total number of pages: 159

Remarks:

Order of re-scan issued on .....